

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2004, 15:25:50 ; Search time 60 Seconds  
(without alignments)  
1186.700 Million cell updates/sec

Title: US-09-857-826B-17  
Perfect score: 252  
Sequence: 1 MAELEFVQIIIIIVVMVMV.....PLESAIWSKEDKQKGHPL 252

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0  
Searched: 1586107 seqs, 282547505 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A Geneseq 29Jan04.\*  
1: geneseqp1980s.\*  
2: geneseqp1990s.\*  
3: geneseqp2000s.\*  
4: geneseqp2001s.\*  
5: geneseqp2002s.\*  
6: geneseqp2003as.\*  
7: geneseqp2003bs.\*  
8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	252	100.0	252	3 AAB01388	Aab01388 Neuron-as
2	252	100.0	252	3 AAB18449	Aab18449 A human T
3	252	100.0	252	5 ABG61805	Abg61805 Prostate
4	252	100.0	252	6 ABP97219	Abp97219 Tumour-as
5	252	100.0	252	6 ABP97218	Abp97218 Tumour-as
6	252	100.0	252	7 ADB75589	Adb75589 Prostate
7	252	100.0	252	7 ADC37325	Adc37325 Nuclear f
8	252	100.0	252	7 AAE39589	Aae39589 Human and
9	252	100.0	285	6 ABP75660	Abp75660 Human sec
10	250	99.2	287	5 AAU78231	Aau78231 Human Mec
11	250	99.2	287	6 ABP97234	Abp97234 Tumour-as
12	250	99.2	287	7 ADC37327	Adc37327 Nuclear f
13	249	98.8	249	7 AAE39591	Aae39591 Human and
14	249	98.8	252	3 AAB18461	Aab18461 A human T
15	237	94.0	237	4 AAM38712	Aam38712 Human pol
16	207	82.1	207	4 AAM40498	Aam40498 Human pol
17	204	81.0	252	3 AAB18462	Aab18462 A human T
18	198	78.6	252	3 AAB18463	Aab18463 A human T
19	181	71.8	241	5 ABG77052	Abg77052 Prostate
20	106	42.1	270	2 AAY74136	Aay74136 Human pro
21	51	20.2	217	3 AAB18466	Aab18466 A murine
22	51	20.2	217	3 AAB18464	Aab18464 A murine
23	51	20.2	217	3 AAB18465	Aab18465 A murine
24	51	20.2	217	3 AAB18450	Aab18450 A murine
25	51	20.2	274	5 AAU78236	Aau78236 Mouse Mec

26	35	13.9	261	4	ABG20316	Abg20316 Novel hum
27	35	13.9	269	3	AAU73422	Aay73422 Human sec
28	35	13.9	269	6	ADA54268	Ada54268 Human pro
29	35	13.9	288	2	AAW60674	Aaw60674 Manic-dep
30	35	13.9	288	7	ADC37453	Adc37453 Nuclear f
31	35	13.9	306	2	AAW60673	Aaw60673 Manic-dep
32	35	13.9	306	7	ADE59611	Ade59611 Human pro
33	35	13.9	316	4	ABG13118	Abg13118 Novel hum
34	35	13.9	334	4	ABG13117	Abg13117 Novel hum
35	35	13.9	475	4	ABG28310	Abg28310 Novel hum
36	35	13.9	485	4	ABG13126	Abg13126 Novel hum
37	20	7.9	204	4	ABG13125	Abg13125 Novel hum
38	20	7.9	235	4	ABG13123	Abg13123 Novel hum
39	14	5.6	14	4	AAU00597	Aam00597 Human pro
40	9	3.6	19	4	ABG20315	Abg20315 Novel hum
41	8	3.2	308	4	ABB71424	Abb71424 Drosophil
42	8	3.2	308	4	ABB66489	Abb66489 Drosophil
43	8	3.2	308	4	ABB66488	Abb66488 Drosophil
44	7	2.8	14	4	AAU00598	Aam00598 Human pro
45	7	2.8	32	4	AAU15296	Aam15296 Peptide #

ALIGNMENTS

RESULT 1  
AAB01388  
ID AAB01388 standard; protein; 252 AA.  
XX  
AC AAB01388;  
XX  
DT 20-OCT-2000 (first entry)  
XX  
DE Neuron-associated protein.  
XX  
KW Neuron associated protein; NEUAP; neurological disorder; epilepsy;  
KW ischemic cerebrovascular disease; stroke; cerebral neoplasm;  
KW Alzheimer's disease; Pick's disease; Huntington's disease; dementia;  
KW Parkinson's disease; demyelinating disease; meningitis; prion disease;  
KW kuru; Creutzfeldt-Jakob disease; neurofibromatosis; cerebral palsy;  
KW muscular dystrophy; central nervous system; CNS;  
KW peripheral nervous system; PNS; myopathy; schizophrenia;  
KW actinic keratosis; arteriosclerosis; atherosclerosis; bursitis;  
KW cirrhosis; hepatitis; mixed connective tissue disease; MCTD;  
KW myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;  
KW autoimmune disease; inflammation; acquired immunodeficiency syndrome;  
KW AIDS; Addison's disease; adult respiratory distress syndrome; allergy;  
KW ankylosing spondylitis; amyloidosis; anaemia; asthma;  
KW Werner syndrome, trauma; human.

Self

FT Modified-site 164 /note= "Potential phosphorylation site"  
FT Modified-site 196 /note= "Potential phosphorylation site"  
FT Modified-site 219 /note= "Potential phosphorylation site"  
FT Modified-site 219 /note= "Potential phosphorylation site"  
XX  
PN WO200034477-A2.  
XX  
PD 15-JUN-2000.  
XX  
XX 10-DEC-1999; 99WO-US030408.  
XX  
XX 11-DEC-1998; 98US-00210083.  
PR 09-FEB-1999; 99US-0119365P.  
PR 16-MAR-1999; 99US-0124687P.  
XX  
XX (INCY-) INCYTE PHARM INC.  
XX  
PI Tang YT, Yue H, Baughn MR, Hillman JL, Lal P, Au-Young J;  
PI Yang J, Lu DAM, Azimzai Y;  
XX  
DR WPI; 2000-423423/36.  
DR N-PSDB; AAA47429.  
XX  
XX New human neuron-associated proteins and polynucleotides encoding them,  
PT useful for diagnosis, treatment and prevention of cell proliferative  
PT disorders including cancer, neuronal and neurological disorders.  
XX  
PS Claim 1; Page 113-114; 145pp; English.  
XX  
CC Human neuron-associated proteins (NEUAP) can be used for for treating or  
CC preventing a disorder associated with decreased expression or activity of  
CC NEUAP. Antagonists of NEUAP are useful for treating or preventing  
CC disorder associated with increased expression or activity of NEUAP. NEUAP  
CC or their fragments or derivatives are useful for treating neurological  
CC disorder such as epilepsy, ischemic cerebrovascular disease, stroke,  
CC cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's  
CC disease, dementia and Parkinson's disease. NEUAPs are also useful for  
CC treating other demyelinating diseases, bacterial and viral meningitis,  
CC prion diseases including kuru, Creutzfeldt-Jakob disease, nutritional and  
CC metabolic diseases of the nervous system, neurofibromatosis, other  
CC developmental disorders of the central nervous system, cerebral palsy,  
CC neuroskeletal disorders, autonomic nervous system disorders, cranial  
CC nerve disorders, spinal cord diseases, muscular dystrophy and other  
CC neuromuscular disorders, peripheral nervous system disorders, inherited,  
CC metabolic, endocrine, and toxic myopathies, mental disorders including  
CC mood, anxiety and schizophrenic disorders, a cell proliferative disorder  
CC such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis,  
CC cirrhosis, hepatitis, mixed connective tissue disease (MCTD),  
CC myelofibrosis, paroxysmal nocturnal haemoglobinuria, cancers of the  
CC adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, and an  
CC autoimmune/inflammatory disorder such as acquired immunodeficiency  
CC syndrome (AIDS), Addison's disease, adult respiratory distress syndrome,  
CC allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, Werner  
CC syndrome, complications of cancer, hemodialysis, and extracorporeal  
CC circulation, viral, bacterial, fungal parasitic, protozoal, and  
CC helminthic infections, and trauma. This protein was given the Incyte ID  
CC no. 1871288CD1  
XX  
SQ Sequence 252 AA;

Query Match 100.0%; Score 252; DB 3; Length 252;  
Best Local Similarity 100.0%; Pred. No. 5.8e-233;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAELEFVQIIIIIVVMVMVWVITCLLSHYKLSARSFISRHSGRRRDLSSGCLWPS 60  
DB 1 MAELEFVQIIIIIVVMVMVWVITCLLSHYKLSARSFISRHSGRRRDLSSGCLWPS 60  
QY 61 ESTVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120  
DB 61 ESTVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120

QY 121 BEPPPYQGCTQLQRLDPEQQLNRESVRAPPNRTIFDSLDMSARLGGPCPPSSNSGIS 180  
DB 121 BEPPPYQGCTQLQRLDPEQQLNRESVRAPPNRTIFDSLDMSARLGGPCPPSSNSGIS 180  
QY 181 ATCYGSGGRMEGPPPTTYSEVIGHYPCSSFOHQSSGPPPSLLEGLRLHHTHIAPLESAAIW 240  
DB 181 ATCYGSGGRMEGPPPTTYSEVIGHYPCSSFOHQSSGPPPSLLEGLRLHHTHIAPLESAAIW 240  
QY 241 SKEKDKQKGHPL 252  
DB 241 SKEKDKQKGHPL 252  
RESULT 2  
AAB18449  
ID AAB18449 standard; protein; 252 AA.  
XX  
AC AAB18449;  
DT 15-JAN-2001 (first entry)  
XX  
DE A human TANGO 261 polypeptide.  
XX  
KW TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;  
KW cellular proliferation; cellular differentiation; cellular adhesion;  
KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
KW intestinal disorder; spleen associated disease; renal disorder;  
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
KW brain herniation; iatrogenic disease; inflammation; meningitis;  
KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;  
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..28 /note= "signal peptide"  
FT Protein 29..253 /note= "mature protein"  
FT  
FT  
XX WO200052022-A1.  
XX  
PD 08-SEP-2000.  
XX  
PF 01-MAR-2000; 2000WO-US005226.  
XX  
PR 01-MAR-1999; 99US-0122458P.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;  
XX  
XX WPI; 2000-579269/54.  
DR N-PSDB; AAA75151.  
XX  
PT Novel human and murine secreted proteins designated TANGO 216, 261, 262,  
PT 266 and 267 useful as modulating agents of cellular processes, e.g. for  
PT treating cancer.  
XX  
PS Claim 2; Fig 5; 175pp; English.  
XX  
CC The present sequence represents a human TANGO 261 polypeptide. The  
CC specification also describes TANGO 266, TANGO 216, TANGO 262, and TANGO  
CC 267. The TANGO polypeptides can be used to modulate cellular  
CC proliferation, modulate cellular differentiation and/or modulate cellular  
CC adhesion. The proteins can be used to treat any von Willebrand factor-  
CC associated disorder, regulate extracellular matrix structuring, cellular  
CC adhesion, and cell trafficking and/or migration, modulate cellular  
CC interactions, modulate cell adhesion in proliferative disorders, such as  
CC cancer, modulate the proliferation, differentiation, and/or function of

CC cells that appear in the bone marrow, and leukocytes, treat bone marrow,  
CC blood and hematopoietic associated diseases and disorders, atelectasis,  
CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial  
CC asthma and bronchiectasis, intestinal disorders, spleen associated  
CC diseases, modulate renal disorders, treat cardiovascular disorders such  
CC as ischemic heart disease, modulate the proliferation, differentiation,  
CC and/or function of bone and cartilage cells and to treat bone and/or  
CC cartilage associated diseases or disorder. They may also be used to treat  
CC disorders associated with the ovaries, cerebral oedema, hydrocephalus,  
CC brain herniations, iatrogenic disease, inflammations, bacterial and viral  
CC meningitis, Alzheimer's Disease, cerebral toxoplasmosis, Parkinson's  
CC disease, multiple sclerosis, brain cancers, hydrocephalus and  
CC encephalitis, and treat hepatic disorders

XX Sequence 252 AA;

Query Match 100.0%; Score 252; DB 3; Length 252;  
Best Local Similarity 100.0%; Pred. No. 5.8e-233;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAELEFVQIIIIIVVMVMVVITCLLSHYKLSARSPISRHSQGRRRDALSSGCLWPS 60  
Db 1 MAELEFVQIIIIIVVMVMVVITCLLSHYKLSARSPISRHSQGRRRDALSSGCLWPS 60  
Qy 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120  
Db 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120  
Qy 121 BEPPYQGPCTQLQRLDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180  
Db 121 BEPPYQGPCTQLQRLDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180  
Qy 181 ATCYSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Db 181 ATCYSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Qy 241 SKEKDQKQGHPL 252  
Db 241 SKEKDQKQGHPL 252

RESULT 3

ABG61805  
ID ABG61805 standard; protein; 252 AA.

XX AC ABG61805;

XX 15-AUG-2002 (first entry)

DT Prostate cancer-associated protein #6.

DE Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.

XX OS Mammalia.

XX PN WO200230268-A2.

XX PD 18-APR-2002.

XX PF 12-OCT-2001; 2001WO-US032045.

XX 13-OCT-2000; 2000US-00687576.

PR 08-DEC-2000; 2000US-00733288.

PR 08-DEC-2000; 2000US-00733742.

PR 24-JAN-2001; 2001US-0263957P.

PR 16-MAR-2001; 2001US-0276791P.

PR 16-MAR-2001; 2001US-0276888P.

PR 06-APR-2001; 2001US-0281922P.

PR 24-APR-2001; 2001US-0286214P.

PR 30-APR-2001; 2001US-00847046.

PR 04-MAY-2001; 2001US-0288589P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;  
XX WI; 2002-471335/50.  
DR N-PSDB; ABK92120.

XX Detecting a prostate cancer-associated transcript in a cell in a patient,  
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,  
PT by determining if prostate cancer-associated genes are expressed in a  
PT prostate tissue.

XX Claim 27; Page 305; 436pp; English.

XX The present invention relates to methods of detecting a prostate cancer-  
CC associated transcript in a cell from a patient. The method comprises  
CC contacting a biological sample from the patient with prostate cancer-  
CC associated polynucleotides (designated PC genes) that selectively  
CC hybridise to a sequence that is at least 80% identical to them. The  
CC prostate cancer-associated polynucleotide sequences are differentially  
CC expressed in prostate tumour tissue or in prostate cancer and are derived  
CC from the tissues of various organisms such as humans or other mammals  
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for  
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-  
CC associated genes are useful for diagnosing or treating prostate cancer,  
CC as well as for identifying modulators of prostate cancer or agents that  
CC inhibit prostate cancer. The nucleic acid sequences are particularly  
CC useful in gene therapy, as a vaccine or in antisense applications.  
CC ABG61800-ABG61944 represent prostate cancer-associated proteins

XX Sequence 252 AA;

Query Match 100.0%; Score 252; DB 5; Length 252;  
Best Local Similarity 100.0%; Pred. No. 5.8e-233;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAELEFVQIIIIIVVMVMVVITCLLSHYKLSARSPISRHSQGRRRDALSSGCLWPS 60  
Db 1 MAELEFVQIIIIIVVMVMVVITCLLSHYKLSARSPISRHSQGRRRDALSSGCLWPS 60  
Qy 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120  
Db 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120  
Qy 121 BEPPYQGPCTQLQRLDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180  
Db 121 BEPPYQGPCTQLQRLDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180  
Qy 181 ATCYSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Db 181 ATCYSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Qy 241 SKEKDQKQGHPL 252  
Db 241 SKEKDQKQGHPL 252

RESULT 4

ABP97219

ID ABP97219 standard; protein; 252 AA.

XX AC ABP97219;

XX 01-JUL-2003 (first entry)

DT Tumour-associated antigenic target protein TAT375 SEQ ID NO:101.

DE Human; tumour-associated antigenic target; TAT; tumour; diagnosis;

XX KW cancer.

XX OS Homo sapiens.

XX PN WO2003024392-A2.



PD 27-MAR-2003.  
XX 11-SEP-2002; 2002WO-US028859.  
PF 18-SEP-2001; 2001US-0323268P.  
PR 19-OCT-2001; 2001US-0339227P.  
PR 07-NOV-2001; 2001US-0336827P.  
PR 20-NOV-2001; 2001US-0331906P.  
PR 02-JAN-2002; 2002US-0345444P.  
PR 03-APR-2002; 2002US-0369724P.  
PR 19-AUG-2002; 2002US-0404809P.  
XX (GETH ) GENENTECH INC.  
XX Frantz G, Hillan KJ, Phillips HS, Polakis P, Spencer SD;  
PI Williams PM, Wu TD, Zhang Z;  
XX WPI; 2003-354551/33.  
DR N-PSDB; ACC49537.  
XX New antibodies against tumor-associated antigenic target polypeptide,  
PT useful for treating or diagnosing tumors or cancers in mammals, e.g.  
PT prostate cancer, lung cancer, prostate adenocarcinomas or renal cell  
PT carcinomas.  
XX Claim 2; Fig 101; 285pp; English.  
XX ACC49493 to ACC49552 encode the human tumour-associated antigenic target  
CC (TAT) proteins given in ABP97175 to ABP97234. The present invention  
CC describes an isolated antibody that binds to a polypeptide having at  
CC least 80 % sequence identity to any of the 60 150-800 residue amino acid  
CC sequences (S1), given in ABP97175 to ABP97234, comprising (S1), lacking  
CC its associated signal peptide, encoded by any of the 60 2000-3000 base  
CC pair sequences (S2), given in ACC49493 to ACC49552. TAT proteins have  
CC cytostatic activity. The antibody can be used for treating or diagnosing  
CC tumours or cancers in mammals, e.g. prostate cancer, lung cancer, breast  
CC cancer, colon cancer, ovarian cancer, prostate adenocarcinomas, renal  
CC cell carcinomas, or thyroid cancer  
XX  
SQ Sequence 252 AA;  
Query Match 100.0%; Score 252; DB 6; Length 252;  
Best Local Similarity 100.0%; Pred. No. 5.8e-233;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAELEFVQIIIIIVVMVMVWVITCLLSHYKLSARSPISRHSQGRRRREDALSSEGCLWPS 60  
Db 1 MAELEFVQIIIIIVVMVMVWVITCLLSHYKLSARSPISRHSQGRRRREDALSSEGCLWPS 60  
QY 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQERFHRFQPTYPYLQHEIDLPTISLSDG 120  
Db 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQERFHRFQPTYPYLQHEIDLPTISLSDG 120  
QY 121 EEPYPYQGPCTQLQRPDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCPSPSSNSGIS 180  
Db 121 EEPYPYQGPCTQLQRPDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCPSPSSNSGIS 180  
QY 181 ATCYGSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Db 181 ATCYGSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
QY 241 SKEKDKQKGHPL 252  
Db 241 SKEKDKQKGHPL 252  
RESULT 5  
ABP97218 standard; protein; 252 AA.  
XX  
AC ABP97218;  
XX 01-JUL-2003 (first entry)

XX Tumour-associated antigenic target protein TAT180 SEQ ID NO:100.  
DE Human; tumour-associated antigenic target; TAT; tumour; diagnosis;  
XX cancer.  
KW Homo sapiens.  
XX WO2003024392-A2.  
XX 27-MAR-2003.  
PD 11-SEP-2002; 2002WO-US028859.  
PF 18-SEP-2001; 2001US-0323268P.  
XX 19-OCT-2001; 2001US-0339227P.  
PR 07-NOV-2001; 2001US-0336827P.  
PR 20-NOV-2001; 2001US-0331906P.  
PR 02-JAN-2002; 2002US-0345444P.  
PR 03-APR-2002; 2002US-0369724P.  
PR 19-AUG-2002; 2002US-0404809P.  
XX (GETH ) GENENTECH INC.  
XX Frantz G, Hillan KJ, Phillips HS, Polakis P, Spencer SD;  
PI Williams PM, Wu TD, Zhang Z;  
XX WPI; 2003-354551/33.  
DR N-PSDB; ACC49536.  
XX New antibodies against tumor-associated antigenic target polypeptide,  
PT useful for treating or diagnosing tumors or cancers in mammals, e.g.  
PT prostate cancer, lung cancer, prostate adenocarcinomas or renal cell  
PT carcinomas.  
XX Claim 2; Fig 100; 285pp; English.  
XX ACC49493 to ACC49552 encode the human tumour-associated antigenic target  
CC (TAT) proteins given in ABP97175 to ABP97234. The present invention  
CC describes an isolated antibody that binds to a polypeptide having at  
CC least 80 % sequence identity to any of the 60 150-800 residue amino acid  
CC sequences (S1), given in ABP97175 to ABP97234, comprising (S1), lacking  
CC its associated signal peptide, encoded by any of the 60 2000-3000 base  
CC pair sequences (S2), given in ACC49493 to ACC49552. TAT proteins have  
CC cytostatic activity. The antibody can be used for treating or diagnosing  
CC tumours or cancers in mammals, e.g. prostate cancer, lung cancer, breast  
CC cancer, colon cancer, ovarian cancer, prostate adenocarcinomas, renal  
CC cell carcinomas, or thyroid cancer  
XX  
SQ Sequence 252 AA;  
Query Match 100.0%; Score 252; DB 6; Length 252;  
Best Local Similarity 100.0%; Pred. No. 5.8e-233;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAELEFVQIIIIIVVMVMVWVITCLLSHYKLSARSPISRHSQGRRRREDALSSEGCLWPS 60  
Db 1 MAELEFVQIIIIIVVMVMVWVITCLLSHYKLSARSPISRHSQGRRRREDALSSEGCLWPS 60  
QY 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQERFHRFQPTYPYLQHEIDLPTISLSDG 120  
Db 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQERFHRFQPTYPYLQHEIDLPTISLSDG 120  
QY 121 EEPYPYQGPCTQLQRPDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCPSPSSNSGIS 180  
Db 121 EEPYPYQGPCTQLQRPDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCPSPSSNSGIS 180  
QY 181 ATCYGSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Db 181 ATCYGSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
QY 241 SKEKDKQKGHPL 252  
Db 241 SKEKDKQKGHPL 252



Db 241 SKEKDQKQGHPL 252

RESULT 6

ADB75589

ID ADB75589 standard; protein; 252 AA.

XX

AC ADB75589;

XX

DT 04-DEC-2003 (first entry)

XX

DE Prostate cancer marker protein.

XX

KW Prostate; cancer; cytostatic; gene therapy; marker.

XX

OS Homo sapiens.

XX

PN WO2003009814-A2.

XX

PD 06-FEB-2003.

XX

PF 25-JUL-2002; 2002WO-US023913.

XX

PR 25-JUL-2001; 2001US-0307982P.

PR 22-AUG-2001; 2001US-0314356P.

PR 25-SEP-2001; 2001US-0325020P.

PR 12-DEC-2001; 2001US-0341746P.

PR 05-MAR-2002; 2002US-0362158P.

XX

PA (MILL-) MILLENNIUM PHARM INC.

XX

PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;

PI Hoersh S, Kamatkar S, Wonsey AM, Glatk K, Zhao X, Anderson D;

XX

DR WPI; 2003-248033/24.

XX

XX New nucleic acid molecule, useful for diagnosing or treating prostate cancer.

PT

PS Disclosure; SEQ ID NO 413; 99pp; English.

XX

CC The invention relates to newly discovered cancer markers associated with the cancerous state of prostate cells. Also disclosed is a method of assessing whether a patient is afflicted with prostate cancer. The method of the invention involves assessing whether a patient is afflicted with prostate cancer by comparing the level of expression of a marker in a patient sample and the normal level of expression of the marker in a control non-prostate cancer sample, where a significant increase in the level of expression of the marker in the patient sample and the normal level indicates that the patient is afflicted with prostate cancer.

CC Nucleic acids of the invention are useful for diagnosing or treating prostate cancer, and may be useful in gene therapy. Sequences given in ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 252 AA;

Query Match 100.0%; Score 252; DB 7; Length 252;

Best Local Similarity 100.0%; Pred. No. 5.8e-233;

Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAELFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGQRRREDALSSEGCLWPS 60

Db 1 MAELFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGQRRREDALSSEGCLWPS 60

Qy 61 ESTVSGNGIPEQVYAPRPTDRLAVPPFAQRFHFRFOPTYPYLQHEIDLPTISLSDG 120

Db 61 ESTVSGNGIPEQVYAPRPTDRLAVPPFAQRFHFRFOPTYPYLQHEIDLPTISLSDG 120

Qy 121 EEPYPYQGPCTQLRDPQQLELNRESVRAPPNRTIFDSDLMDSARLGGPCPPSSNSGIS 180

Db 121 EEPYPYQGPCTQLRDPQQLELNRESVRAPPNRTIFDSDLMDSARLGGPCPPSSNSGIS 180

Qy 181 ATCYGSGGRMEGPPPTTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHTHIAPLESAAIW 240

Db 181 ATCYGSGGRMEGPPPTTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHTHIAPLESAAIW 240

Qy 241 SKEKDQKQGHPL 252

Db 241 SKEKDQKQGHPL 252

RESULT 7

ADC37325

ID ADC37325 standard; protein; 252 AA.

XX

AC ADC37325;

XX

DT 18-DEC-2003 (first entry)

XX

DE Nuclear factor kappa B (NF-kappaB) activating protein, SEQ ID 158.

XX

KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease; cancer; infectious disease; bone disease; AIDS;

KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;

KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;

XX Neuroprotective; Nootropic; Cardiant; Gene therapy; human.

OS Homo sapiens.

XX

PN WO2003048202-A2.

XX

PD 12-JUN-2003.

XX

PF 03-DEC-2002; 2002WO-JP012644.

XX

PR 03-DEC-2001; 2001JP-00368692.

PR 05-DEC-2001; 2001US-0335829P.

PR 03-OCT-2002; 2002JP-00291302.

PR 04-OCT-2002; 2002US-0415769P.

XX

PA (ASAH) ASAHI KASEI KK.

XX

PI Matsuda A, Muramatsu S;

XX

DR WPI; 2003-505282/47.

DR N-PSDB; ADC37324.

XX

PT New purified protein that activates nuclear factor kappa B (NF-kappaB), useful for treating inflammation, autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS, neurodegenerative diseases or ischemic disorders.

PT

XX Claim 1; SEQ ID NO 158; 938pp; English.

PS The present invention relates to novel proteins and their coding sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-kappaB). The proteins and their coding sequences are useful for treating a disease associated with NF-kappaB activation, such as inflammation, autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS, neurodegenerative diseases, or ischaemic disorders.

CC

XX Sequence 252 AA;

Query Match 100.0%; Score 252; DB 7; Length 252;

Best Local Similarity 100.0%; Pred. No. 5.8e-233;

Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAELFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGQRRREDALSSEGCLWPS 60

Db 1 MAELFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGQRRREDALSSEGCLWPS 60

Qy 61 ESTVSGNGIPEQVYAPRPTDRLAVPPFAQRFHFRFOPTYPYLQHEIDLPTISLSDG 120

Db 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120  
Qy 121 BEPPYQGPCTQLRDPQQLELNRESVRAPPNRTIFDSDLMD SARLGPCPPSSNSGIS 180  
Db 121 BEPPYQGPCTQLRDPQQLELNRESVRAPPNRTIFDSDLMD SARLGPCPPSSNSGIS 180  
Qy 181 ATCYSGGRMEGPPPTYSYEVIGHYPCSSFHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Db 181 ATCYSGGRMEGPPPTYSYEVIGHYPCSSFHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Qy 241 SKEKDQKQGHPL 252  
Db 241 SKEKDQKQGHPL 252

RESULT 8  
AAE39589  
ID AAE39589 standard; protein; 252 AA.  
XX  
AC AAE39589;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Human androgen-regulated gene (ARG), PMEPA1 protein.  
XX  
KW Human; androgen-regulated gene; ARG; PMEPA1; prostate cancer;  
KW chromosome 20q13.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 9..25  
FT /note= "Transmembrane domain"  
XX  
PN US6566130-B1.  
XX  
PD 20-MAY-2003.  
XX  
PF 26-JAN-2001; 2001US-00769482.  
XX  
PR 28-JAN-2000; 2000US-0178772P.  
PR 31-JAN-2000; 2000US-0179045P.  
XX  
PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.  
XX  
PI Srivastava S, Moul JW, Xu LL, Segawa T;  
XX  
DR WPI; 2003-719644/68.  
DR N-PSDB; AAD60105.  
XX  
PT Novel isolated androgen-regulated gene designated as PMEPA1 useful for  
PT selecting primers and probes for detecting prostate cancer cells in  
PT biological samples by nucleic acid amplification techniques.  
XX  
PS Claim 1; Col 29-30; 58pp; English.  
XX  
CC The invention relates to an isolated androgen-regulated gene (ARG)  
CC designated as PMEPA1. The invention is useful for selecting primers and  
CC probes for detecting prostate cancer cells in a biological sample by  
CC using nucleic acid amplification techniques. The present sequence is  
CC human PMEPA1 protein. PMEPA1 gene chromosome 20q13  
SQ Sequence 252 AA;

Query Match 100.0%; Score 252; DB 7; Length 252;  
Best Local Similarity 100.0%; Pred. No. 5.8e-233;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAELEFVQIIIIIVVMVMVVTCLLSHYKLSARSFISHSQGRREDALSSGCLMPS 60  
Db 1 MAELEFVQIIIIIVVMVMVVTCLLSHYKLSARSFISHSQGRREDALSSGCLMPS 60  
Qy 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120

Db 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120  
Qy 121 BEPPYQGPCTQLRDPQQLELNRESVRAPPNRTIFDSDLMD SARLGPCPPSSNSGIS 180  
Db 121 BEPPYQGPCTQLRDPQQLELNRESVRAPPNRTIFDSDLMD SARLGPCPPSSNSGIS 180  
Qy 181 ATCYSGGRMEGPPPTYSYEVIGHYPCSSFHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Db 181 ATCYSGGRMEGPPPTYSYEVIGHYPCSSFHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Qy 241 SKEKDQKQGHPL 252  
Db 241 SKEKDQKQGHPL 252

RESULT 9  
ABP75660  
ID ABP75660 standard; protein; 285 AA.  
XX  
AC ABP75660;  
XX  
DT 10-FEB-2003 (first entry)  
XX  
DE Human secretory polypeptide SPTM SEQ ID NO 844.  
XX  
KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;  
KW asthma; Crohn's disease; neurological disorder; epilepsy; cancer;  
KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;  
KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;  
KW anti-inflammatory; immunosuppressive; neuroprotective; neurotropic;  
KW neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;  
KW antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;  
KW secretory polynucleotide; secretory protein.  
XX  
OS Homo sapiens.  
XX  
PN WO200283876-A2.  
XX  
PD 24-OCT-2002.  
XX  
PF 27-MAR-2002; 2002WO-US009921.  
XX  
PR 29-MAR-2001; 2001US-0280067P.  
PR 29-MAR-2001; 2001US-0280068P.  
PR 16-MAY-2001; 2001US-0291280P.  
PR 17-MAY-2001; 2001US-0291829P.  
PR 17-MAY-2001; 2001US-0291849P.  
PR 19-JUN-2001; 2001US-0299428P.  
PR 20-JUN-2001; 2001US-0299776P.  
PR 20-JUN-2001; 2001US-030001P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;  
PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Anshey SR;  
PI Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstein EH;  
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;  
PI Flores V, Marwaha R, Lo A, Lan RY, Urashka MB;  
XX  
DR WPI; 2003-075543/07.  
DR N-PSDB; ABZ36103.  
XX  
PT New human secretory proteins and polynucleotides, useful for diagnosing,  
PT treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),  
PT neurological disorders (e.g. Alzheimer's), or cell proliferations or  
PT cancers.  
XX  
PS Claim 27; SEQ ID NO 844; 458pp + Sequence Listing; English.  
XX  
CC The invention relates to a secretory polynucleotide (designated sptm)  
CC comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a  
CC naturally occurring polynucleotide sequence at least 90 % identical to

CC the polynucleotide sequence, a polynucleotide complementary to them or an  
CC RNA equivalent of them. The polypeptide or polynucleotide are useful for  
CC treating, preventing or diagnosing a disease or condition associated with  
CC the expression of functional SPTM. These are particularly useful for  
CC diagnosing, treating or preventing autoimmune/inflammatory disorders  
CC (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's  
CC disease), neurological disorders (e.g. epilepsy, Huntington's disease,  
CC dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,  
CC multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,  
CC schizophrenia or amnesia), or cell proliferative disorders (e.g.  
CC psoriasis, polycythemia vera, or cancers including adenocarcinoma,  
CC leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,  
CC breast, cervix or prostate). The present sequence is one of the SPTM  
CC proteins of the invention (ABP75384-ABP75962). Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX

SQ Sequence 285 AA;  
Query Match 100.0%; Score 252; DB 6; Length 285;  
Best Local Similarity 100.0%; Pred. No. 6.4e-233;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPS 60  
DB 34 MALEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPS 93  
QY 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120  
DB 94 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 153  
QY 121 BEPPYQGPCTLQLRDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNGIS 180  
DB 154 BEPPYQGPCTLQLRDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNGIS 213  
QY 181 ATCYSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLGTRLHHTHIAPLESAAIW 240  
DB 214 ATCYSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLGTRLHHTHIAPLESAAIW 273  
QY 241 SKEKDQKQGHPL 252  
DB 274 SKEKDQKQGHPL 285

RESULT 10  
AAU78231  
ID AAU78231 standard; protein; 287 AA.  
XX  
AC AAU78231;  
XX  
DT 05-JUN-2002 (first entry)  
XX  
DE Human Mechanically Induced Vascular Receptor 1, MIVR-1.  
XX  
KW Human; MIVR-1; Mechanically Induced Vascular; receptor; cytostatic;  
KW cardiant; cerebroprotective; antiarteriosclerotic; cardiac cell;  
KW anti-apoptotic; vascular endothelial cell; cardiac hypertrophy;  
KW myocardial infarction; stroke; arteriosclerosis; heart failure.  
XX  
OS Homo sapiens.  
XX  
FN WO200216416-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 21-AUG-2001; 2001WO-US026089.  
XX  
PR 22-AUG-2000; 2000US-0227159P.  
XX  
PA (BGM ) BRIGHAM & WOMENS HOSPITAL INC.  
PA (PFIZ ) PFIZER INC.  
XX

PI Lee RT, Landschulz KT, Kennedy SP, Thompson JF, Turi TG;  
XX WPI; 2002-280912/32.  
DR N-PSDB; ABK12137.  
XX  
PT Novel nucleic acid molecule encoding Mechanically Induced Vascular  
PT Receptor-1 polypeptide, useful for treating cardiovascular diseases.  
XX  
PS Claim 14; Page 88-89; 105pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule encoding a  
CC Mechanically Induced Vascular Receptor (MIVR)-1 polypeptide having  
CC cardiac cell anti-apoptotic activity and fragments of it provided they  
CC are not identical to Genbank sequences AI761441.1, AI594390, NM\_004338  
CC and AQ177461. Also included are expression vectors, host cells, the MIVR-  
CC 1 polypeptide, MIVR-1 binding peptides, modulators of MIVR-1, contacting  
CC a molecule having cardiac cell anti-apoptotic activity with a candidate  
CC agent, where the molecule is a nucleic acid molecule comprising MIVR-1,  
CC IEX-1, VDU-1, BTG-2 and TIS-11d or its expression product, determining  
CC if the anti-apoptotic activity is modulated and thereby identifying a  
CC modulator. The cardiac cell anti-apoptotic molecules and nucleic acids  
CC of the invention are useful for treating, diagnosing and monitoring  
CC progression of such diseases and disorders as characterised by increased  
CC apoptotic cell-death of vascular endothelial cells e.g. cardiac  
CC hypertrophy, myocardial infarction, stroke, arteriosclerosis and heart  
CC failure. The present sequence represents human MIVR-1  
XX  
SQ Sequence 287 AA;

Query Match 99.2%; Score 250; DB 5; Length 287;  
Best Local Similarity 100.0%; Pred. No. 5.3e-231;  
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPS 62  
DB 38 ELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPS 97  
QY 63 TVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 122  
DB 98 TVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 157  
QY 123 PPPYQGPCTLQLRDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNGISAT 182  
DB 158 PPPYQGPCTLQLRDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNGISAT 217  
QY 183 CYGSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLGTRLHHTHIAPLESAAIWSK 242  
DB 218 CYGSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLGTRLHHTHIAPLESAAIWSK 277  
QY 243 EKDKQKQGHPL 252  
DB 278 EKDKQKQGHPL 287

RESULT 11  
ABP97234  
ID ABP97234 standard; protein; 287 AA.  
XX  
AC ABP97234;  
XX  
DT 01-JUL-2003 (first entry)  
XX  
DE Tumour-associated antigenic target protein TAT379 cDNA SEQ ID NO:120.  
XX  
KW Human; tumour-associated antigenic target; TAT; tumour; diagnosis;  
KW cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO2003024392-A2.  
XX  
PD 27-MAR-2003.  
XX



PF 11-SEP-2002; 2002WO-US028859.  
 XX  
 PR 18-SEP-2001; 2001US-0323268P.  
 PR 19-OCT-2001; 2001US-0339227P.  
 PR 07-NOV-2001; 2001US-0336827P.  
 PR 20-NOV-2001; 2001US-0331906P.  
 PR 02-JAN-2002; 2002US-0345444P.  
 PR 03-APR-2002; 2002US-0369724P.  
 PR 19-AUG-2002; 2002US-0404809P.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Frantz G, Hillan KJ, Phillips HS, Polakis P, Spencer SD;  
 PI Williams PM, Wu TD, Zhang Z;  
 XX  
 DR WPI; 2003-354551/33.  
 DR N-PSDB; ACC49552.  
 XX  
 PT New antibodies against tumor-associated antigenic target polypeptide,  
 PT useful for treating or diagnosing tumors or cancers in mammals, e.g.  
 PT prostate cancer, lung cancer, prostate adenocarcinomas or renal cell  
 PT carcinomas.  
 XX  
 PS Claim 2; Fig 120; 285pp; English.  
 XX  
 CC ACC49493 to ACC49552 encode the human tumour-associated antigenic target  
 CC (TAT) proteins given in ABP97175 to ABP97234. The present invention  
 CC describes an isolated antibody that binds to a polypeptide having at  
 CC least 80 % sequence identity to any of the 60 150-800 residue amino acid  
 CC sequences (S1), given in ABP97175 to ABP97234, comprising (S1), lacking  
 CC its associated signal peptide, encoded by any of the 60 2000-3000 base  
 CC pair sequences (S2), given in ACC49493 to ACC49552. TAT proteins have  
 CC cytostatic activity. The antibody can be used for treating or diagnosing  
 CC tumours or cancers in mammals, e.g. prostate cancer, lung cancer, breast  
 CC cancer, colon cancer, ovarian cancer, prostate adenocarcinomas, renal  
 CC cell carcinomas, or thyroid cancer  
 XX  
 SQ Sequence 287 AA;

Query Match 99.2%; Score 250; DB 6; Length 287;  
 Best Local Similarity 100.0%; Pred. No. 5.3e-231;  
 Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ELEFVQIIIIIVVMVMVWVITCLSHYKLSARSFISRHSGRRRREDALSSGCLWPSES 62  
 DB 38 ELEFVQIIIIIVVMVMVWVITCLSHYKLSARSFISRHSGRRRREDALSSGCLWPSES 97  
 QY 63 TVSGNGIPEPQVYAPPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDGEE 122  
 DB 98 TVSGNGIPEPQVYAPPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDGEE 157  
 QY 123 PPHYQGPCTQLRDPEQQLNRESVRAPNRTIFDSDLMD SARLGCGCPSSNSGISAT 182  
 DB 158 PPHYQGPCTQLRDPEQQLNRESVRAPNRTIFDSDLMD SARLGCGCPSSNSGISAT 217  
 QY 183 CYGSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSK 242  
 DB 218 CYGSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSK 277  
 QY 243 EKDKQKGHPL 252  
 DB 278 EKDKQKGHPL 287

RESULT 12  
 ADC37327  
 ID ADC37327 standard; protein; 287 AA.

XX  
 AC ADC37327;  
 XX  
 DT 18-DEC-2003 (first entry)  
 XX  
 DE Nuclear factor kappa B (NF-kappaB) activating protein, SEQ ID 160.

XX  
 KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;  
 KW cancer; infectious disease; bone disease; AIDS;  
 KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;  
 KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;  
 KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003048202-A2.  
 XX  
 PD 12-JUN-2003.  
 XX  
 PF 03-DEC-2002; 2002WO-JP012644.  
 XX  
 PR 03-DEC-2001; 2001JP-00368692.  
 PR 05-DEC-2001; 2001US-0335829P.  
 PR 03-OCT-2002; 2002JP-00291302.  
 PR 04-OCT-2002; 2002US-0415769P.  
 XX  
 PA (ASAH ) ASAH KASEI KK.  
 XX  
 PI Matsuda A, Muramatsu S;  
 XX  
 DR WPI; 2003-505282/47.  
 DR N-PSDB; ADC37326.  
 XX  
 PT New purified protein that activates nuclear factor kappa B (NF-kappaB),  
 PT useful for treating inflammation, autoimmune diseases, cancers,  
 PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or  
 PT ischemic disorders.  
 XX  
 PS Claim 1; SEQ ID NO 160; 938pp; English.  
 XX  
 CC The present invention relates to novel proteins and their coding  
 CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-  
 CC kappaB). The proteins and their coding sequences are useful for treating  
 CC a disease associated with NF-kappaB activation, such as inflammation,  
 CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,  
 CC neurodegenerative diseases, or ischaemic disorders.  
 XX  
 SQ Sequence 287 AA;

Query Match 99.2%; Score 250; DB 7; Length 287;  
 Best Local Similarity 100.0%; Pred. No. 5.3e-231;  
 Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ELEFVQIIIIIVVMVMVWVITCLSHYKLSARSFISRHSGRRRREDALSSGCLWPSES 62  
 DB 38 ELEFVQIIIIIVVMVMVWVITCLSHYKLSARSFISRHSGRRRREDALSSGCLWPSES 97  
 QY 63 TVSGNGIPEPQVYAPPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDGEE 122  
 DB 98 TVSGNGIPEPQVYAPPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDGEE 157  
 QY 123 PPHYQGPCTQLRDPEQQLNRESVRAPNRTIFDSDLMD SARLGCGCPSSNSGISAT 182  
 DB 158 PPHYQGPCTQLRDPEQQLNRESVRAPNRTIFDSDLMD SARLGCGCPSSNSGISAT 217  
 QY 183 CYGSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSK 242  
 DB 218 CYGSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSK 277  
 QY 243 EKDKQKGHPL 252  
 DB 278 EKDKQKGHPL 287

RESULT 13  
 AAE39591  
 ID AAE39591 standard; protein; 249 AA.  
 XX  
 AC AAE39591;

XX DT 18-DEC-2003 (first entry)  
XX DE Human androgen-regulated gene (ARG), PMEPA1 protein fragment.  
XX KW Human; androgen-regulated gene; ARG; PMEPA1; prostate cancer; chromosome 20q13.  
XX OS Homo sapiens.  
XX PN US6566130-B1.  
XX PD 20-MAY-2003.  
XX PF 26-JAN-2001; 2001US-00769482.  
XX PR 28-JAN-2000; 2000US-0178772P.  
XX PR 31-JAN-2000; 2000US-0179045P.  
XX PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.  
XX PI Srivastava S, Moul JW, Xu LL, Segawa T;  
XX DR WPI; 2003-719644/68.  
XX PT Novel isolated androgen-regulated gene designated as PMEPA1 useful for selecting primers and probes for detecting prostate cancer cells in biological samples by nucleic acid amplification techniques.  
XX PS Example 7; Col 31-32; 58pp; English.  
XX CC The invention relates to an isolated androgen-regulated gene (ARG) designated as PMEPA1. The invention is useful for selecting primers and probes for detecting prostate cancer cells in a biological sample by using nucleic acid amplification techniques. The present sequence is human PMEPA1 protein fragment. PMEPA1 gene chromosome 20q13  
XX CC  
XX SQ Sequence 249 AA;  
Query Match 98.8%; Score 249; DB 7; Length 249;  
Best Local Similarity 100.0%; Pred. No. 4.3e-230;  
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 AELEFVQIIIVVMVMVVITCLLSHYKLSARSPISRHSQGRREDALSSEGCLWPSE 61  
DB 1 AELEFVQIIIVVMVMVVITCLLSHYKLSARSPISRHSQGRREDALSSEGCLWPSE 60  
QY 62 STVSGNGIPEPQVYAPPRPTDRLAVPPPAQRERFHRFQPTYPYLQHEIDLPTISLSDGE 121  
DB 61 STVSGNGIPEPQVYAPPRPTDRLAVPPPAQRERFHRFQPTYPYLQHEIDLPTISLSDGE 120  
QY 122 EPPYQGPCTQLRDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGISA 181  
DB 121 EPPYQGPCTQLRDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGISA 180  
QY 182 TCYSGGRMEGPPPTYSEVIGHYPCSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIWS 241  
DB 181 TCYSGGRMEGPPPTYSEVIGHYPCSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIWS 240  
QY 242 KEKDQKQGH 250  
DB 241 KEKDQKQGH 249

RESULT 14  
AAB18461  
ID AAB18461 standard; protein; 252 AA.  
XX AC AAB18461;  
XX 15-JAN-2001 (first entry)  
DT  
XX A human TANGO 261 polypeptide clone.

XX KW TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;  
KW cellular proliferation; cellular differentiation; cellular adhesion;  
KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
KW intestinal disorder; spleen associated disease; renal disorder;  
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
KW brain herniation; iatrogenic disease; inflammation; meningitis;  
KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;  
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder.  
XX OS Homo sapiens.  
XX PN WO200052022-A1.  
XX PD 08-SEP-2000.  
XX PF 01-MAR-2000; 2000WO-US005226.  
XX PR 01-MAR-1999; 99US-0122458P.  
XX PA (MILL-) MILLENNIUM PHARM INC.  
PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;  
XX WPI; 2000-579269/54.  
DR N-PSDB; AAA75163.  
XX PT Novel human and murine secreted proteins designated TANGO 216, 261, 262,  
PT 266 and 267 useful as modulating agents of cellular processes, e.g. for  
PT treating cancer.  
XX PT Disclosure; Page; 175pp; English.  
CC AAB18461-63 represent human TANGO 261 proteins. The specification also  
CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO  
CC polypeptides can be used to modulate cellular proliferation, modulate  
CC cellular differentiation and/or modulate cellular adhesion. The proteins  
CC can be used to treat any von Willebrand factor-associated disorder, and  
CC regulate extracellular matrix structuring, cellular adhesion, and cell  
CC trafficking and/or migration, modulate cellular interactions, modulate  
CC cell adhesion in proliferative disorders, such as cancer, modulate the  
CC proliferation, differentiation, and/or function of cells that appear in  
CC the bone marrow, and leukocytes, treat bone marrow, blood and  
CC hematopoietic associated diseases and disorders, atelectasis, pulmonary  
CC congestion or oedema, emphysema, chronic bronchitis, bronchial asthma and  
CC bronchiectasis, intestinal disorders, spleen associated diseases,  
CC modulate renal disorders, treat cardiovascular disorders such as ischemic  
CC heart disease, modulate the proliferation, differentiation, and/or  
CC function of bone and cartilage cells and to treat bone and/or cartilage  
CC associated diseases or disorder. They may also be used to treat disorders  
CC associated with the ovaries, and cerebral oedema, hydrocephalus, brain  
CC herniations, iatrogenic disease, inflammations, bacterial and viral  
CC meningitis, Alzheimer's Disease, cerebral toxoplasmosis, Parkinson's  
CC disease, multiple sclerosis, brain cancers, hydrocephalus and  
CC encephalitis, and treat hepatic disorders. note: the present sequence  
CC does not appear in the specification; it was created using information  
CC provided  
XX SQ Sequence 252 AA;

Query Match 98.8%; Score 249; DB 3; Length 252;  
Best Local Similarity 100.0%; Pred. No. 4.3e-230;  
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 LEFVQIIIVVMVMVVITCLLSHYKLSARSPISRHSQGRREDALSSEGCLWPSEST 63  
DB 4 LEFVQIIIVVMVMVVITCLLSHYKLSARSPISRHSQGRREDALSSEGCLWPSEST 63  
QY 64 VSGNGIPEPQVYAPPRPTDRLAVPPPAQRERFHRFQPTYPYLQHEIDLPTISLSDGEEP 123  
DB 64 VSGNGIPEPQVYAPPRPTDRLAVPPPAQRERFHRFQPTYPYLQHEIDLPTISLSDGEEP 123

QY 124 PPYQGPCTQLRDEQQLNELNRESVRAPPNRTIFDSDLMDSARLGPCPPSSNSGISATC 183  
|||  
Db 124 PPYQGPCTQLRDEQQLNELNRESVRAPPNRTIFDSDLMDSARLGPCPPSSNSGISATC 183  
QY 184 YGSGGRMEGPPPTSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIWSKE 243  
|||  
Db 184 YGSGGRMEGPPPTSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIWSKE 243  
QY 244 KDKQKGHP 252  
|||  
Db 244 KDKQKGHP 252

RESULT 15  
AAM38712  
ID AAM38712 standard; protein; 237 AA.  
XX  
AC AAM38712;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 1857.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00489725.

XX 25-APR-2000; 2000US-00552317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00662191.

XX 19-OCT-2000; 2000US-00693036.

XX 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;

XX Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX N-PSDB; AAI57868.

XX Novel nucleic acids and polypeptides, useful for treating disorders such

XX as central nervous system injuries.

XX Example 3; SEQ ID NO 1857; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the

XX encoded polypeptides (AAM38642-AAM42213) with nootropic,

XX immunosuppressant and cytostatic activity. The polynucleotides are useful

XX in gene therapy. A composition containing a polypeptide or polynucleotide

XX of the invention may be used to treat diseases of the peripheral nervous

XX system, such as peripheral nervous injuries, peripheral neuropathy and

XX localised neuropathies and central nervous system diseases, such as

XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

XX utilisation of the activities such as: Immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
CC part of the printed specification

XX Sequence 237 AA;

SQ Query Match 94.0%; Score 237; DB 4; Length 237;

Best Local Similarity 100.0%; Pred. No. 1.3e-218;

Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MMVMVVVITCLLSHYKLSARSPISRHSQRRRREDALSSGCLWPSSTVSGNGIPEPQVY 75

Db 1 MMVMVVVITCLLSHYKLSARSPISRHSQRRRREDALSSGCLWPSSTVSGNGIPEPQVY 60

QY 76 APPRPTDRLAVPPAQRERFQPTYPYLQHEIDLPTTISLSDGEEPPPYQGCTQLQR 135

Db 61 APPRPTDRLAVPPAQRERFQPTYPYLQHEIDLPTTISLSDGEEPPPYQGCTQLQR 120

QY 136 DPEQOLELNRESVRAPPNRTIFDSDLMDSARLGPCPPSSNSGISATCYGSGRMEGPPPP 195

Db 121 DPEQOLELNRESVRAPPNRTIFDSDLMDSARLGPCPPSSNSGISATCYGSGRMEGPPPP 180

QY 196 TYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIWSKEKDKQKGHP 252

Db 181 TYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIWSKEKDKQKGHP 237

Search completed: May 20, 2004, 15:34:47

Job time : 63 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2004, 15:33:51 ; Search time 23 Seconds  
(without alignments)  
565.641 Million cell updates/sec

Title: US-09-857-826B-17  
Perfect score: 252  
Sequence: 1 MAELEFVQIIIIIVVMVMV.....PLSAAIWSKXKQKHPL 252

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents\_AA\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	252	100.0	252	4	US-09-769-482-3
2	249	98.8	249	4	US-09-769-482-11
3	35	13.9	244	4	US-09-769-482-12
4	35	13.9	288	4	US-09-091-952A-4
5	35	13.9	306	4	US-09-091-952A-3
6	14	5.6	69	4	US-09-621-976-6022
7	7	2.8	18	4	US-09-091-952A-5
8	7	2.8	269	4	US-09-134-000C-5214
9	7	2.8	272	4	US-09-252-991A-23410
10	7	2.8	285	4	US-09-149-476-470
11	7	2.8	289	3	US-08-872-979-6
12	7	2.8	326	4	US-09-107-532A-3665
13	7	2.8	369	4	US-09-252-991A-27769
14	7	2.8	407	4	US-09-252-991A-20436
15	7	2.8	448	4	US-09-518-657-2
16	7	2.8	450	4	US-09-518-657-4
17	7	2.8	540	4	US-09-513-057C-33
18	7	2.8	551	4	US-09-252-991A-26416
19	7	2.8	589	1	US-08-399-646-2
20	7	2.8	589	1	US-08-607-321-2
21	7	2.8	589	2	US-08-961-240-2
22	7	2.8	589	2	US-08-605-501-2
23	7	2.8	589	1	US-08-399-646-12
24	7	2.8	596	1	US-08-607-321-12
25	7	2.8	596	2	US-08-961-240-12
26	7	2.8	596	2	US-08-605-501-12
27	7	2.8	664	2	US-08-852-153-8

28	7	2.8	666	4	US-09-228-986-68	Sequence 68, Appl
29	6	2.4	21	4	US-09-288-143-178	Sequence 178, App
30	6	2.4	37	3	US-08-470-369-19	Sequence 19, Appl
31	6	2.4	37	6	5304637-19	Patent No. 5304637
32	6	2.4	46	3	US-08-905-223-347	Sequence 347, App
33	6	2.4	74	4	US-09-621-976-7191	Sequence 7191, Ap
34	6	2.4	86	1	US-08-390-858B-39	Sequence 39, Appl
35	6	2.4	95	4	US-09-309-572-8	Sequence 8, Appli
36	6	2.4	95	4	US-09-718-096-8	Sequence 8, Appli
37	6	2.4	99	4	US-09-621-976-4793	Sequence 4793, Ap
38	6	2.4	101	4	US-09-621-976-7124	Sequence 7124, Ap
39	6	2.4	102	4	US-09-289-198-308	Sequence 308, App
40	6	2.4	102	4	US-09-429-755-308	Sequence 308, App
41	6	2.4	103	1	US-08-376-296-10	Sequence 10, Appl
42	6	2.4	110	1	US-08-844-064-8	Sequence 8, Appli
43	6	2.4	110	3	US-09-009-433-8	Sequence 8, Appli
44	6	2.4	112	4	US-09-621-976-6250	Sequence 6250, Ap
45	6	2.4	120	4	US-09-615-192A-393	Sequence 393, App

ALIGNMENTS

RESULT 1  
US-09-769-482-3  
; Sequence 3, Application US/09769482  
; Patent No. 6566130  
; GENERAL INFORMATION:  
; APPLICANT: SRIVASTAVA, SHIV  
; APPLICANT: MOUL, JUDD W.  
; APPLICANT: XU, LINDA L.  
; APPLICANT: SEGAWA, TAKEHIKO  
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED  
; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY  
; FILE REFERENCE: 04995.0057-00000  
; CURRENT APPLICATION NUMBER: US/09/769,482  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,772  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/179,045  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 252  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-769-482-3

Query Match 100.0%; Score 252; DB 4; Length 252;  
Best Local Similarity 100.0%; Pred. No. 2.4e-239;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAELEFVQIIIIIVVMVMVMVVITCLLSHYKLSARSFISRHSGQRRREDALSSGCLWPS	60
Db	1	MAELEFVQIIIIIVVMVMVMVVITCLLSHYKLSARSFISRHSGQRRREDALSSGCLWPS	60
Qy	61	ESTVSGNGIPEPQVYAPPRPTDRLAVPPFAQRERFHFQPTYPYLQHEIDLPTISLSDG	120
Db	61	ESTVSGNGIPEPQVYAPPRPTDRLAVPPFAQRERFHFQPTYPYLQHEIDLPTISLSDG	120
Qy	121	EEPPPYQGPGCTQLQRDPQQLNRESVRAPPNRTIFDSILMDSARLGGPCPPSSNSGIS	180
Db	121	EEPPPYQGPGCTQLQRDPQQLNRESVRAPPNRTIFDSILMDSARLGGPCPPSSNSGIS	180
Qy	181	ATCYGSGRMEGPPPTYSVIGHYPGSSFOHQSSGGPPSLLEGTRLLHHTHIAPLESAAIW	240
Db	181	ATCYGSGRMEGPPPTYSVIGHYPGSSFOHQSSGGPPSLLEGTRLLHHTHIAPLESAAIW	240
Qy	241	SKKDKQKQKHPL	252
Db	241	SKKDKQKQKHPL	252

filed 1-26-07  
earliest 1-28-00

```
RESULT 2
US-09-769-482-11
; Sequence 11, Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-769-482-11

Query Match      98.8%; Score 249; DB 4; Length 249;
Best Local Similarity 100.0%; Pred. No. 2.1e-236;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRRREDALSSGCLWPSE 61
      |||||||
DB      1 AELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRRREDALSSGCLWPSE 60

QY      62 STVSGNGIPEQVYAPRPTDRLAVPPFAQERFHRFQPTYPYLQHEIDLPTISLSDGE 121
      |||||||
DB      61 STVSGNGIPEQVYAPRPTDRLAVPPFAQERFHRFQPTYPYLQHEIDLPTISLSDGE 120

QY      122 EPPYQGPCTQLRDPEQOLELNRESVRAPNRTIFDSLDMSARLGCGPCPPSSNGISA 181
      |||||||
DB      121 EPPYQGPCTQLRDPEQOLELNRESVRAPNRTIFDSLDMSARLGCGPCPPSSNGISA 180

QY      182 TCYSGGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIWS 241
      |||||||
DB      181 TCYSGGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIWS 240

QY      242 KEKQKQKH 250
      |||||||
DB      241 KEKQKQKH 249

RESULT 3
US-09-769-482-12
; Sequence 12, Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
```

```
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-769-482-12

Query Match      13.9%; Score 35; DB 4; Length 244;
Best Local Similarity 100.0%; Pred. No. 2e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      106 QHEIDLPTISLSDGEEPPPYQGPPCTQLQRDEQQ 140
      |||||||
DB      103 QHEIDLPTISLSDGEEPPPYQGPPCTQLQRDEQQ 137

RESULT 4
US-09-091-952A-4
; Sequence 4, Application US/09091952A
; Patent No. 6458532
; GENERAL INFORMATION:
; APPLICANT: Detera-Wadleigh, Sevilla D.
; APPLICANT: Gershon, Elliot S.
; APPLICANT: Badner, Judith A.
; APPLICANT: Goldin, Lynn R.
; APPLICANT: Berrettini, Wade H.
; APPLICANT: Yoshikawa, Takeo
; APPLICANT: Sanders, Alan R.
; APPLICANT: Esterling, Lisa E.
; TITLE OF INVENTION: Chromosomal Markers and Diagnostic
; Tests for Manic-Depressive Illness
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/091,952A
; FILING DATE: 19-Apr-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,278
; FILING DATE: 28-OCT-1996
; APPLICATION NUMBER: PCT/US97/19381
; FILING DATE: 28-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 015280-297100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 288 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..288
; OTHER INFORMATION: Clone 22 isoform 2
; alternatively spliced
; protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
```

US-09-091-952A-4

Query Match 13.9%; Score 35; DB 4; Length 288;  
Best Local Similarity 100.0%; Pred. No. 2.4e-26;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 QHEIDLPTTISLSDGEEPPPYQGPCTQLQLRDPEQQ 140  
Db 145 QHEIDLPTTISLSDGEEPPPYQGPCTQLQLRDPEQQ 179

RESULT 5

US-09-091-952A-3  
; Sequence 3, Application US/09091952A  
; Patent No. 6458532

GENERAL INFORMATION:

; APPLICANT: Detera-Wadleigh, Sevilla D.  
; Gershon, Elliot S.  
; Badner, Judith A.  
; Goldin, Lynn R.  
; Berrettini, Wade H.  
; Yoshikawa, Takeo  
; Sanders, Alan R.  
; Esterling, Lisa E.

; TITLE OF INVENTION: Chromosomal Markers and Diagnostic Tests for Manic-Depressive Illness

; NUMBER OF SEQUENCES: 197

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/091,952A

; FILING DATE: 19-Apr-1999

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/029,278

; FILING DATE: 28-OCT-1996

; APPLICATION NUMBER: PCT/US97/19381

; FILING DATE: 28-OCT-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Smith, Timothy L.

; REGISTRATION NUMBER: 35,367

; REFERENCE/DOCKET NUMBER: 015280-297100US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300

; TELEX: <Unknown>

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 306 amino acids

; TYPE: amino acid

; STRANDEDNESS: <Unknown>

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FEATURE:

; NAME/KEY: mat\_peptide

; LOCATION: 1...306

; OTHER INFORMATION: Clone 22 isoform 1,

; unspliced protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-091-952A-3

Query Match 13.9%; Score 35; DB 4; Length 306;  
Best Local Similarity 100.0%; Pred. No. 2.5e-26;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 QHEIDLPTTISLSDGEEPPPYQGPCTQLQLRDPEQQ 140  
Db 163 QHEIDLPTTISLSDGEEPPPYQGPCTQLQLRDPEQQ 197

RESULT 6

US-09-621-976-6022  
; Sequence 6022, Application US/09621976  
; Patent No. 6639063

GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 6022  
; LENGTH: 69  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-09-621-976-6022

Query Match 5.6%; Score 14; DB 4; Length 69;  
Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 GGPCPPSSNSGISA 181  
Db 4 GGPCPPSSNSGISA 17

RESULT 7

US-09-091-952A-5  
; Sequence 5, Application US/09091952A  
; Patent No. 6458532

GENERAL INFORMATION:

; APPLICANT: Detera-Wadleigh, Sevilla D.  
; Gershon, Elliot S.  
; Badner, Judith A.  
; Goldin, Lynn R.  
; Berrettini, Wade H.  
; Yoshikawa, Takeo  
; Sanders, Alan R.  
; Esterling, Lisa E.

; TITLE OF INVENTION: Chromosomal Markers and Diagnostic Tests for Manic-Depressive Illness

; NUMBER OF SEQUENCES: 197

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/091,952A

; FILING DATE: 19-Apr-1999

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/029,278

; FILING DATE: 28-OCT-1996

; APPLICATION NUMBER: PCT/US97/19381

; FILING DATE: 28-OCT-1997



ATTORNEY/AGENT INFORMATION:  
 NAME: Smith, Timothy L.  
 REGISTRATION NUMBER: 35,367  
 REFERENCE/DOCKET NUMBER: 015280-297100US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 TELEX: <Unknown>  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 18 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: <Unknown>  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 FRAGMENT TYPE: internal  
 FEATURE:  
 NAME/KEY: -  
 LOCATION: 1...18  
 OTHER INFORMATION: alternatively spliced portion  
 lacking from isoform 2  
 SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
 US-09-091-952A-5

Query Match 2.8%; Score 7; DB 4; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 5.6;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 54 EGCLWPS 60  
 |||||  
 Db 1 EGCLWPS 7

RESULT 8

US-09-134-000C-5214  
 Sequence 5214, Application US/09134000C  
 Patent No. 6617156  
 GENERAL INFORMATION:  
 APPLICANT: Lynn Doucette-Stamm et al  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 032796-032  
 CURRENT APPLICATION NUMBER: US/09/134,000C  
 CURRENT FILING DATE: 1998-08-13  
 PRIOR APPLICATION NUMBER: US 60/055,778  
 PRIOR FILING DATE: 1997-08-15  
 NUMBER OF SEQ ID NOS: 6812  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 5214  
 LENGTH: 269  
 TYPE: PRT  
 ORGANISM: Enterococcus faecalis  
 US-09-134-000C-5214

Query Match 2.8%; Score 7; DB 4; Length 269;  
 Best Local Similarity 100.0%; Pred. No. 67;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 168 GGPCPPS 174  
 |||||  
 Db 84 GGPCPPS 90

RESULT 9

US-09-252-991A-23410  
 Sequence 23410, Application US/09252991A  
 Patent No. 6551795  
 GENERAL INFORMATION:  
 APPLICANT: Marc J. Rubenfield et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 107196.136  
 CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18  
 PRIOR APPLICATION NUMBER: US 60/074,788  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 60/094,190  
 PRIOR FILING DATE: 1998-07-27  
 NUMBER OF SEQ ID NOS: 33142  
 SEQ ID NO 23410  
 LENGTH: 272  
 TYPE: PRT  
 ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-23410

Query Match 2.8%; Score 7; DB 4; Length 272;  
 Best Local Similarity 100.0%; Pred. No. 68;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 166 RLGGPCP 172  
 |||||  
 Db 33 RLGGPCP 39

RESULT 10

US-09-149-476-470  
 Sequence 470, Application US/09149476  
 Patent No. 6420526  
 GENERAL INFORMATION:  
 APPLICANT: Rosen et al.  
 TITLE OF INVENTION: 186 Human Secreted proteins  
 FILE REFERENCE: P2002P1  
 CURRENT APPLICATION NUMBER: US/09/149,476  
 CURRENT FILING DATE: 1998-09-08  
 EARLIER APPLICATION NUMBER: PCT/US98/04493  
 EARLIER FILING DATE: 1998-03-06  
 EARLIER APPLICATION NUMBER: 60/040,162  
 EARLIER FILING DATE: 1997-03-07  
 EARLIER APPLICATION NUMBER: 60/040,333  
 EARLIER FILING DATE: 1997-03-07  
 EARLIER APPLICATION NUMBER: 60/038,621  
 EARLIER FILING DATE: 1997-03-07  
 EARLIER APPLICATION NUMBER: 60/040,626  
 EARLIER FILING DATE: 1997-03-07  
 EARLIER APPLICATION NUMBER: 60/040,334  
 EARLIER FILING DATE: 1997-03-07  
 EARLIER APPLICATION NUMBER: 60/040,336  
 EARLIER FILING DATE: 1997-03-07  
 EARLIER APPLICATION NUMBER: 60/040,163  
 EARLIER FILING DATE: 1997-03-07  
 EARLIER APPLICATION NUMBER: 60/047,600  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,615  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,597  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,502  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,633  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,583  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,617  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,618  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,503  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,592  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,581  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,584  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,500  
 EARLIER FILING DATE: 1997-05-23

;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/056,911
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/056,636
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/056,874
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/056,910
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/056,864
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/056,631
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/056,845
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/056,892
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/057,761
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/047,595
;	EARLIER	FILING DATE:	1997-05-23
;	EARLIER	APPLICATION NUMBER:	60/047,599
;	EARLIER	FILING DATE:	1997-05-23
;	EARLIER	APPLICATION NUMBER:	60/047,588
;	EARLIER	FILING DATE:	1997-05-23
;	EARLIER	APPLICATION NUMBER:	60/047,585
;	EARLIER	FILING DATE:	1997-05-23
;	EARLIER	APPLICATION NUMBER:	60/047,586
;	EARLIER	FILING DATE:	1997-05-23
;	EARLIER	APPLICATION NUMBER:	60/047,590
;	EARLIER	FILING DATE:	1997-05-23
;	EARLIER	APPLICATION NUMBER:	60/047,594
;	EARLIER	FILING DATE:	1997-05-23
;	EARLIER	APPLICATION NUMBER:	60/047,589
;	EARLIER	FILING DATE:	1997-05-23
;	EARLIER	APPLICATION NUMBER:	60/047,593
;	EARLIER	FILING DATE:	1997-05-23
;	EARLIER	APPLICATION NUMBER:	60/047,614
;	EARLIER	FILING DATE:	1997-05-23
;	EARLIER	APPLICATION NUMBER:	60/043,578
;	EARLIER	FILING DATE:	1997-04-11
;	EARLIER	APPLICATION NUMBER:	60/043,576
;	EARLIER	FILING DATE:	1997-04-11
;	EARLIER	APPLICATION NUMBER:	60/047,501
;	EARLIER	FILING DATE:	1997-05-23
;	EARLIER	APPLICATION NUMBER:	60/043,670
;	EARLIER	FILING DATE:	1997-04-11
;	EARLIER	APPLICATION NUMBER:	60/056,632
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/056,664
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/056,876
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/056,881
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/056,909
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/056,875
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/056,862
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/056,887
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/056,908
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/048,964
;	EARLIER	FILING DATE:	1997-06-06
;	EARLIER	APPLICATION NUMBER:	60/057,650
;	EARLIER	FILING DATE:	1997-09-05
;	EARLIER	APPLICATION NUMBER:	60/056,884
;	EARLIER	FILING DATE:	1997-08-22
;	EARLIER	APPLICATION NUMBER:	60/057,669
;	EARLIER	FILING DATE:	1997-09-05

; EARLIER APPLICATION NUMBER: 60/049,610  
 ; EARLIER FILING DATE: 1997-06-13  
 ; EARLIER APPLICATION NUMBER: 60/061,060  
 ; EARLIER FILING DATE: 1997-10-02

Query Match 2.8%; Score 7; DB 4; Length 285;  
 Best Local Similarity 100.0%; Pred. No. 71;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 214 SSGPSSL 220  
 Db 38 SSGPSSL 44

RESULT 11

US-08-872-979-6  
 ; Sequence 6, Application US/08872979  
 ; Patent No. 6074844  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hillman, Jennifer L.  
 ; APPLICANT: Lal, Preeti  
 ; APPLICANT: Corley, Neil C.  
 ; TITLE OF INVENTION: TWO NEW HUMAN MEMBRANE FUSION PROTEINS  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: PastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/872,979  
 FILING DATE: Herewith  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:

ATTORNEY/AGENT INFORMATION:  
 NAME: Billings, Lucy J.  
 REGISTRATION NUMBER: 36,749  
 REFERENCE/DOCKET NUMBER: PP-0320 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-855-0555  
 TELEFAX: 415-845-4166  
 TELEX:

INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 289 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: GenBank  
 CLONE: 924268

Query Match 2.8%; Score 7; DB 3; Length 289;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 IIIIVV 15  
 Db 265 IIIIVV 271

RESULT 12

US-09-107-532A-3665

; Sequence 3665, Application US/09107532A  
 ; Patent No. 6583275  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
 ; NUMBER OF SEQUENCES: 7310  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
 ; STREET: 100 Beaver Street  
 ; CITY: Waltham  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02354  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: CD-ROM ISO9660  
 ; COMPUTER: PC  
 ; OPERATING SYSTEM: <Unknown>  
 ; SOFTWARE: ASCII  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/107,532A  
 ; FILING DATE: 30-Jun-1998  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/085,598  
 ; FILING DATE: 14 May 1998  
 ; APPLICATION NUMBER: 60/051571  
 ; FILING DATE: July 2, 1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Ariniello, Pamela Deneke  
 ; REGISTRATION NUMBER: 40,489  
 ; REFERENCE/DOCKET NUMBER: GTC-012  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (781)893-5007  
 ; TELEFAX: (781)893-8277  
 ; INFORMATION FOR SEQ ID NO: 3665:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 326 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: YES  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Enterococcus faecium  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (B) LOCATION 1...326  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 3665:  
 ; US-09-107-532A-3665

Query Match 2.8%; Score 7; DB 4; Length 326;  
 Best Local Similarity 100.0%; Pred. No. 80;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 162 MDSARLG 168  
 Db 93 MDSARLG 99

RESULT 13

US-09-252-991A-27769  
 ; Sequence 27769, Application US/09252991A  
 ; Patent No. 6551795

; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27



; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 27769  
; LENGTH: 369  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27769

Query Match 2.8%; Score 7; DB 4; Length 369;  
Best Local Similarity 100.0%; Pred. No. 90;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 HSQGRRR 47  
|||||||  
Db 258 HSQGRRR 264

RESULT 14  
US-09-252-991A-20436  
; Sequence 20436, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 20436  
; LENGTH: 407  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-20436

Query Match 2.8%; Score 7; DB 4; Length 407;  
Best Local Similarity 100.0%; Pred. No. 98;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 78 PRPTDRL 84  
|||||||  
Db 246 PRPTDRL 252

RESULT 15  
US-09-518-657-2  
; Sequence 2, Application US/09518657  
; Patent No. 6335188  
; GENERAL INFORMATION:  
; APPLICANT: Schardl, Christopher L.  
; APPLICANT: Wang, Jinghong  
; TITLE OF INVENTION: Endophyte Ergot Alkaloid Synthetic Compounds, Compounds  
; TITLE OF INVENTION: Which Encode Therefor and Related Methods  
; FILE REFERENCE: P-1060  
; CURRENT APPLICATION NUMBER: US/09/518,657  
; CURRENT FILING DATE: 2000-03-03  
; EARLIER APPLICATION NUMBER: 60/125,490  
; EARLIER FILING DATE: 1999-03-22  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 448  
; TYPE: PRT  
; ORGANISM: Neotyphodium coenophialum  
US-09-518-657-2

Query Match 2.8%; Score 7; DB 4; Length 448;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 IPEPOVY 75  
|||||||  
Db 346 IPEPOVY 352  
  
Search completed: May 20, 2004, 15:37:22  
Job time : 24 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2004, 15:36:21 ; Search time 48 Seconds  
(without alignments)  
1464.339 Million cell updates/sec

Title: US-09-857-826B-17  
Perfect score: 252  
Sequence: 1 MALEFVQIIIVVMVMV.....PLESAIWSKERKQKHPL 252

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1149313 segs, 278921704 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 1149313

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:\*

1:	/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17:	/cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	252	100.0	252	10	US-09-796-753-56
2	252	100.0	252	10	US-09-821-812-3
3	252	100.0	252	14	US-10-205-823-413
4	252	100.0	252	14	US-10-241-220-100
5	252	100.0	252	14	US-10-241-220-101
6	252	100.0	252	14	US-10-301-822-209
7	252	100.0	252	14	US-10-390-045-3
8	252	100.0	252	16	US-10-434-479-3
9	250	99.2	287	9	US-09-934-249-2
10	250	99.2	287	14	US-10-241-220-120
11	250	99.2	287	15	US-10-295-027-128
12	249	98.8	249	14	US-10-390-045-11
13	249	98.8	249	16	US-10-434-479-11
14	181	71.8	241	14	US-10-000-256A-158
15	51	20.2	217	10	US-09-796-753-58

16	51	20.2	243	16	US-10-434-479-68	Sequence 68, Appl
17	51	20.2	274	9	US-09-934-249-13	Sequence 13, Appl
18	35	13.9	244	14	US-10-390-045-12	Sequence 12, Appl
19	35	13.9	244	16	US-10-434-479-12	Sequence 12, Appl
20	35	13.9	269	15	US-10-094-749-1836	Sequence 1836, Ap
21	35	13.9	288	14	US-10-251-598-4	Sequence 0, Appli
22	35	13.9	306	12	US-10-087-192-1668	Sequence 1668, Ap
23	35	13.9	306	14	US-10-251-598-3	Sequence 0, Appli
24	8	3.2	453	12	US-10-425-114-49175	Sequence 49175, A
25	8	3.2	585	12	US-10-425-114-71739	Sequence 71739, A
26	7	2.8	18	14	US-10-251-598-5	Sequence 5, Appli
27	7	2.8	32	9	US-09-864-761-35019	Sequence 35019, A
28	7	2.8	37	9	US-09-864-761-37387	Sequence 37387, A
29	7	2.8	41	9	US-09-864-761-34567	Sequence 34567, A
30	7	2.8	64	12	US-10-424-599-192751	Sequence 192751,
31	7	2.8	70	12	US-10-424-599-210969	Sequence 210969,
32	7	2.8	72	9	US-09-864-761-45782	Sequence 45782, A
33	7	2.8	88	12	US-10-335-977-8356	Sequence 8356, Ap
34	7	2.8	89	12	US-10-424-599-216094	Sequence 216094,
35	7	2.8	94	12	US-10-335-977-8357	Sequence 8357, Ap
36	7	2.8	105	9	US-09-864-761-34191	Sequence 34191, A
37	7	2.8	109	9	US-09-864-761-33541	Sequence 33541, A
38	7	2.8	119	14	US-10-156-761-14686	Sequence 14686, A
39	7	2.8	129	12	US-10-425-114-48837	Sequence 48837, A
40	7	2.8	136	12	US-10-424-599-206919	Sequence 206919,
41	7	2.8	146	16	US-10-427-692-651	Sequence 651, App
42	7	2.8	158	12	US-10-425-114-37993	Sequence 37993, A
43	7	2.8	205	12	US-10-425-114-46070	Sequence 46070, A
44	7	2.8	228	15	US-10-200-012-44	Sequence 44, Appl
45	7	2.8	247	12	US-10-425-114-53348	Sequence 53348, A

ALIGNMENTS

RESULT 1  
US-09-796-753-56  
; Sequence 56, Application US/09796753  
; Publication No. US20030027998A1  
; GENERAL INFORMATION:  
; APPLICANT: McCarthy, Sean A.  
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF  
; FILE REFERENCE: 7853-227-999  
; CURRENT APPLICATION NUMBER: US/09/796,753  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 09/183,175  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: 09/223,094  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/223,546  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/224,246  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/259,388  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR APPLICATION NUMBER: 60/122,458  
; PRIOR FILING DATE: 1999-03-01  
; PRIOR APPLICATION NUMBER: 09/312,359  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 09/336,536  
; PRIOR FILING DATE: 1999-06-18  
; PRIOR APPLICATION NUMBER: 09/342,687  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: 09/345,464  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: 09/365,164  
; PRIOR FILING DATE: 1999-07-30  
; PRIOR APPLICATION NUMBER: 09/399,723  
; PRIOR FILING DATE: 1999-09-20  
; PRIOR APPLICATION NUMBER: 09/409,634  
; PRIOR FILING DATE: 1999-09-30  
; PRIOR APPLICATION NUMBER: 09/471,179  
; PRIOR FILING DATE: 1999-12-23

```

; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 56
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-796-753-56

```

	Query Match	100.0%;	Score 252;	DB 10;	Length 252;
	Best Local Similarity	100.0%;	Pred. No. 1.7e-231;		
	Matches 252;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAELEFVQIIIIVVVMVMVVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPS	60		
Db	1	MAELEFVQIIIIVVVMVMVVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPS	60		
QY	61	ESTVSGNGIPEPQVYAPPRTDRLAVPPFAQRERFHRPQTPPYLQHEIDLPTTISLSDG	120		
Db	61	ESTVSGNGIPEPQVYAPPRTDRLAVPPFAQRERFHRPQTPPYLQHEIDLPTTISLSDG	120		
QY	121	BEPPPYQGCPCTLQLRDEPQQLNRESVRAPPNRTIFDSLDMSARLGGPCPPSSNSGIS	180		
Db	121	BEPPPYQGCPCTLQLRDEPQQLNRESVRAPPNRTIFDSLDMSARLGGPCPPSSNSGIS	180		
QY	181	ATCYGSGGRMEGPPPTYSEVIGHYPGSSFOHQOSSGPPSLLEGTRLNHTHIAPLESAAIW	240		
Db	181	ATCYGSGGRMEGPPPTYSEVIGHYPGSSFOHQOSSGPPSLLEGTRLNHTHIAPLESAAIW	240		
QY	241	SKEKOKQKGHPL	252		
Db	241	SKEKOKQKGHPL	252		

```

RESULT 2
US-09-821-812-3
; Sequence 3, Application US/09821812
; Publication No. US20030166520A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Biaoyang
; TITLE OF INVENTION: Androgen Regulated Prostate Specific
; TITLE OF INVENTION: Nucleic Acids
; FILE REFERENCE: P-IS 4373
; CURRENT APPLICATION NUMBER: US/09/821,812
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens

```

```
US-09-821-812-3

Query Match      100.0%; Score 252; DB 10; Length 252;
Best Local Similarity 100.0%; Pred. NO. 1.7e-231;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY    1 MAELEFVQIIIIVVMMVMVVITCLLSHYKLSARSPISRHSGRRREDALSSGGLWPS 60
      |||||||
Db     1 MAELEFVQIIIIVVMMVMVVITCLLSHYKLSARSPISRHSGRRREDALSSGGLWPS 60

QY    61 ESTVSGNGIPEPQVIAPPRPTDRLAVPPFAQRERFHRFQTPPYLQHEIDLPTTISLDG 120
      |||||||
Db     61 ESTVSGNGIPEPQVIAPPRPTDRLAVPPFAQRERFHRFQTPPYLQHEIDLPTTISLDG 120

QY    121 BEPPPYQGPCCTLQLRDPEQQLELNRESVRAPPNRTIFDSDLMDSARLGGPCPPSSNSGIS 180
      |||||||
Db     121 BEPPPYQGPCCTLQLRDPEQQLELNRESVRAPPNRTIFDSDLMDSARLGGPCPPSSNSGIS 180

QY    181 ATCYGSGGRMEGPPTYSEVIGHYPGSSFQHQQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
      |||||||
Db     181 ATCYGSGGRMEGPPTYSEVIGHYPGSSFQHQQSSGPPSLLEGTRLHHTHIAPLESAAIW 240

QY    241 SKKKOKQKGHP L 252
      |||||||
Db     241 SKKKOKQKGHP L 252
```

RESULT 3  
US-10-205-823-413  
; Sequence 413, Application US/10205823  
; Publication No. US20030108963A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John B.  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Gannavarapu, Manjula  
; APPLICANT: Gorbacheva, Bella  
; APPLICANT: Hoersch, Sebastian  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Monsey, Angela M.  
; APPLICANT: Glatt, Karen  
; APPLICANT: Zhao, Xumei  
; APPLICANT: Anderson, Dustin  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER  
; FILE REFERENCE: MRI-044  
; CURRENT APPLICATION NUMBER: US/10/205,823  
; CURRENT FILING DATE: 2002-07-25  
; PRIOR APPLICATION NUMBER: 60/307,982  
; PRIOR FILING DATE: 2001-07-25  
; PRIOR APPLICATION NUMBER: 60/314,356  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: 60/325,020  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: 60/341,746  
; PRIOR FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: 60/362,158  
; PRIOR FILING DATE: 2002-03-05  
; NUMBER OF SEQ ID NOS: 455  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 413  
; LENGTH: 252  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-205-823-413

```

Query Match          100.0%; Score 252; DB 14; Length 252;
Best Local Similarity 100.0%; Pred. No. 1.7e-231;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



Db 1 MALEFVQIIIIIVVMMVVVITCLLSHYKLSARSPISRHSQRRREDALSSEGCLWPS 60  
QY 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120  
Db 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120  
QY 121 BEPPPYQGCTQLQRLDPEQQLNRESVRAPPNRTIFDSLDMSARLGPCPPSSNSGIS 180  
Db 121 BEPPPYQGCTQLQRLDPEQQLNRESVRAPPNRTIFDSLDMSARLGPCPPSSNSGIS 180  
QY 181 ATCYGSGRMEGPPPTYSEVIGHYPCSSFHQQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Db 181 ATCYGSGRMEGPPPTYSEVIGHYPCSSFHQQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
QY 241 SKEKDKQKGHPL 252  
Db 241 SKEKDKQKGHPL 252

RESULT 4

US-10-241-220-100  
; Sequence 100, Application US/10241220  
; Publication No. US20030148408A1

GENERAL INFORMATION:

; APPLICANT: Frantz, Gretchen  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Phillips, Heidi  
; APPLICANT: Polakis, Paul  
; APPLICANT: Spencer, Susan  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wu, Thomas  
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND

; TITLE OF INVENTION: TREATMENT OF TUMOR

; FILE REFERENCE: P5010R1-US

; CURRENT APPLICATION NUMBER: US/10/241,220

; CURRENT FILING DATE: 2002-12-13

; NUMBER OF SEQ ID NOS: 120

; SEQ ID NO 100

; LENGTH: 252

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-241-220-100

Query Match 100.0%; Score 252; DB 14; Length 252;  
Best Local Similarity 100.0%; Pred. No. 1.7e-231;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEFVQIIIIIVVMMVVVITCLLSHYKLSARSPISRHSQRRREDALSSEGCLWPS 60  
Db 1 MALEFVQIIIIIVVMMVVVITCLLSHYKLSARSPISRHSQRRREDALSSEGCLWPS 60  
QY 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120  
Db 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120  
QY 121 BEPPPYQGCTQLQRLDPEQQLNRESVRAPPNRTIFDSLDMSARLGPCPPSSNSGIS 180  
Db 121 BEPPPYQGCTQLQRLDPEQQLNRESVRAPPNRTIFDSLDMSARLGPCPPSSNSGIS 180  
QY 181 ATCYGSGRMEGPPPTYSEVIGHYPCSSFHQQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Db 181 ATCYGSGRMEGPPPTYSEVIGHYPCSSFHQQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
QY 241 SKEKDKQKGHPL 252  
Db 241 SKEKDKQKGHPL 252

RESULT 5

US-10-241-220-101

; Sequence 101, Application US/10241220

; Publication No. US20030148408A1

; GENERAL INFORMATION:  
; APPLICANT: Frantz, Gretchen  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Phillips, Heidi  
; APPLICANT: Polakis, Paul  
; APPLICANT: Spencer, Susan  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wu, Thomas  
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND

; TITLE OF INVENTION: TREATMENT OF TUMOR

; FILE REFERENCE: P5010R1-US

; CURRENT APPLICATION NUMBER: US/10/241,220

; CURRENT FILING DATE: 2002-12-13

; NUMBER OF SEQ ID NOS: 120

; SEQ ID NO 101

; LENGTH: 252

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-241-220-101

Query Match 100.0%; Score 252; DB 14; Length 252;  
Best Local Similarity 100.0%; Pred. No. 1.7e-231;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEFVQIIIIIVVMMVVVITCLLSHYKLSARSPISRHSQRRREDALSSEGCLWPS 60  
Db 1 MALEFVQIIIIIVVMMVVVITCLLSHYKLSARSPISRHSQRRREDALSSEGCLWPS 60  
QY 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120  
Db 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120  
QY 121 BEPPPYQGCTQLQRLDPEQQLNRESVRAPPNRTIFDSLDMSARLGPCPPSSNSGIS 180  
Db 121 BEPPPYQGCTQLQRLDPEQQLNRESVRAPPNRTIFDSLDMSARLGPCPPSSNSGIS 180  
QY 181 ATCYGSGRMEGPPPTYSEVIGHYPCSSFHQQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Db 181 ATCYGSGRMEGPPPTYSEVIGHYPCSSFHQQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
QY 241 SKEKDKQKGHPL 252  
Db 241 SKEKDKQKGHPL 252

RESULT 6

US-10-301-822-209

; Sequence 209, Application US/10301822

; Publication No. US20030148410A1

GENERAL INFORMATION:

; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Berger, Allison  
; APPLICANT: Guillemette, Tracy L.  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Thibodeau, Stephen N.  
; APPLICANT: Burt, Lawrence J.

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND

; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; TITLE OF INVENTION: THERAPY OF COLON CANCER

; FILE REFERENCE: MP01-029P2RNM

; CURRENT APPLICATION NUMBER: US/10/301,822

; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US 60/339,971

; PRIOR FILING DATE: 2001-12-10

; PRIOR APPLICATION NUMBER: US 60/361,978

; PRIOR FILING DATE: 2002-03-05

; PRIOR APPLICATION NUMBER: US 60/381,988

; PRIOR FILING DATE: 2002-05-20

; NUMBER OF SEQ ID NOS: 228

; SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 209  
LENGTH: 252  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-10-301-822-209

Query Match 100.0%; Score 252; DB 14; Length 252;  
Best Local Similarity 100.0%; Pred. No. 1.7e-231;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALEFVQIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPS 60  
Db 1 MALEFVQIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPS 60  
Qy 61 ESTVSGNGIPEQVYAPRPTDLAVPPPAQRERFHRFQTPYQLQHEIDLPTTISLSDG 120  
Db 61 ESTVSGNGIPEQVYAPRPTDLAVPPPAQRERFHRFQTPYQLQHEIDLPTTISLSDG 120  
Qy 121 EEPYPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMSARLGCGPCPPSSNSGIS 180  
Db 121 EEPYPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMSARLGCGPCPPSSNSGIS 180  
Qy 181 ATCYSGGGRMEGPPPTYSVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Db 181 ATCYSGGGRMEGPPPTYSVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Qy 241 SKEKDKQKGHPL 252  
Db 241 SKEKDKQKGHPL 252

RESULT 7  
US-10-390-045-3  
Sequence 3, Application US/10390045  
Publication No. US20030170713A1  
GENERAL INFORMATION:  
APPLICANT: SRIVASTAVA, SHIV  
APPLICANT: MOUL, JUDD W.  
APPLICANT: SEGAWA, TAKEHIKO  
TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED  
TITLE OF INVENTION: POYNUCLEOTIDE ARRAY  
FILE REFERENCE: 04995.0057-00000  
CURRENT APPLICATION NUMBER: US/10/390,045  
CURRENT FILING DATE: 2003-03-18  
PRIOR APPLICATION NUMBER: US/09/769,482  
PRIOR FILING DATE: 2001-01-26  
PRIOR APPLICATION NUMBER: 60/178,772  
PRIOR FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: 60/179,045  
PRIOR FILING DATE: 2000-01-31  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 252  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-390-045-3

Query Match 100.0%; Score 252; DB 14; Length 252;  
Best Local Similarity 100.0%; Pred. No. 1.7e-231;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALEFVQIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPS 60  
Db 1 MALEFVQIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPS 60  
Qy 61 ESTVSGNGIPEQVYAPRPTDLAVPPPAQRERFHRFQTPYQLQHEIDLPTTISLSDG 120  
Db 61 ESTVSGNGIPEQVYAPRPTDLAVPPPAQRERFHRFQTPYQLQHEIDLPTTISLSDG 120  
Qy 121 EEPYPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMSARLGCGPCPPSSNSGIS 180  
Db 121 EEPYPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMSARLGCGPCPPSSNSGIS 180

Db 121 EEPYPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMSARLGCGPCPPSSNSGIS 180  
Qy 181 ATCYSGGGRMEGPPPTYSVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Db 181 ATCYSGGGRMEGPPPTYSVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Qy 241 SKEKDKQKGHPL 252  
Db 241 SKEKDKQKGHPL 252

RESULT 8  
US-10-434-479-3  
Sequence 3, Application US/10434479  
Publication No. US20040092469A1  
GENERAL INFORMATION:  
APPLICANT: SRIVASTAVA, SHIV  
APPLICANT: MOUL, JUDD W.  
APPLICANT: XU, LINDA L.  
TITLE OF INVENTION: ANDROGEN-REGULATED PMEPA1 GENE AND POLYPEPTIDES  
FILE REFERENCE: 04995.0057-02000  
CURRENT APPLICATION NUMBER: US/10/434,479  
CURRENT FILING DATE: 2003-05-09  
PRIOR APPLICATION NUMBER: 10/390,045  
PRIOR FILING DATE: 2003-03-18  
PRIOR APPLICATION NUMBER: 09/769,482  
PRIOR FILING DATE: 2001-01-26  
PRIOR APPLICATION NUMBER: 60/178,772  
PRIOR FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: 60/179,045  
PRIOR FILING DATE: 2000-01-31  
NUMBER OF SEQ ID NOS: 81  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 252  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-434-479-3

Query Match 100.0%; Score 252; DB 16; Length 252;  
Best Local Similarity 100.0%; Pred. No. 1.7e-231;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALEFVQIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPS 60  
Db 1 MALEFVQIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPS 60  
Qy 61 ESTVSGNGIPEQVYAPRPTDLAVPPPAQRERFHRFQTPYQLQHEIDLPTTISLSDG 120  
Db 61 ESTVSGNGIPEQVYAPRPTDLAVPPPAQRERFHRFQTPYQLQHEIDLPTTISLSDG 120  
Qy 121 EEPYPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMSARLGCGPCPPSSNSGIS 180  
Db 121 EEPYPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMSARLGCGPCPPSSNSGIS 180  
Qy 181 ATCYSGGGRMEGPPPTYSVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Db 181 ATCYSGGGRMEGPPPTYSVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Qy 241 SKEKDKQKGHPL 252  
Db 241 SKEKDKQKGHPL 252

RESULT 9  
US-09-934-249-2  
Sequence 2, Application US/09934249  
Patent No. US20020115081A1  
GENERAL INFORMATION:  
APPLICANT: Lee, Richard T.  
APPLICANT: Landschulz, Katherine T.  
APPLICANT: Turi, Thomas G.  
APPLICANT: Thompson, John F.

```

; APPLICANT: Kennedy, Scott P.
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR CONDITIONS
; FILE REFERENCE: P0738/7001/ERP/KA
; CURRENT APPLICATION NUMBER: US/09/934,249
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,159
; PRIOR FILING DATE: 2000-08-22
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-09-934-249-2

Query Match          99.2%; Score 250; DB 9; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.5e-229;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 ELEFVQIIIIIVMMVMVVITCLLSHYKLSARSFISRHSGQRRREDALSSGCLWPSES 62
Db 38 ELEFVQIIIIIVMMVMVVITCLLSHYKLSARSFISRHSGQRRREDALSSGCLWPSES 97

Qy 63 TVSGNGIPEPQVYAPRPTDRLAVPPFAQRRFHRFQPTYPYLQHEIDLPTTISLSDGEE 122
Db 98 TVSGNGIPEPQVYAPRPTDRLAVPPFAQRRFHRFQPTYPYLQHEIDLPTTISLSDGEE 157

Qy 123 PPHYQGCTLQLRDPQQLELNRESVRAPPNRTIFDSDLMSARLGSGPCPPSSNSGISAT 182
Db 158 PPHYQGCTLQLRDPQQLELNRESVRAPPNRTIFDSDLMSARLGSGPCPPSSNSGISAT 217

Qy 183 CYGSGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSK 242
Db 218 CYGSGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSK 277

Qy 243 EKDKQKGHPL 252
Db 278 EKDKQKGHPL 287

RESULT 10
US-10-241-220-120
; Sequence 120, Application US/10241220
; Publication No. US20030148408A1
; GENERAL INFORMATION:
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Phillips, Heidi
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5010R1-US
; CURRENT APPLICATION NUMBER: US/10/241,220
; CURRENT FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 120
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-241-220-120

Query Match          99.2%; Score 250; DB 14; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.5e-229;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 ELEFVQIIIIIVMMVMVVITCLLSHYKLSARSFISRHSGQRRREDALSSGCLWPSES 62
Db 38 ELEFVQIIIIIVMMVMVVITCLLSHYKLSARSFISRHSGQRRREDALSSGCLWPSES 97

Qy 63 TVSGNGIPEPQVYAPRPTDRLAVPPFAQRRFHRFQPTYPYLQHEIDLPTTISLSDGEE 122
Db 98 TVSGNGIPEPQVYAPRPTDRLAVPPFAQRRFHRFQPTYPYLQHEIDLPTTISLSDGEE 157

Qy 123 PPHYQGCTLQLRDPQQLELNRESVRAPPNRTIFDSDLMSARLGSGPCPPSSNSGISAT 182
Db 158 PPHYQGCTLQLRDPQQLELNRESVRAPPNRTIFDSDLMSARLGSGPCPPSSNSGISAT 217

Qy 183 CYGSGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSK 242
Db 218 CYGSGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSK 277

Qy 243 EKDKQKGHPL 252
Db 278 EKDKQKGHPL 287
```

```

Qy 63 TVSGNGIPEPQVYAPRPTDRLAVPPFAQRRFHRFQPTYPYLQHEIDLPTTISLSDGEE 122
Db 98 TVSGNGIPEPQVYAPRPTDRLAVPPFAQRRFHRFQPTYPYLQHEIDLPTTISLSDGEE 157

Qy 123 PPHYQGCTLQLRDPQQLELNRESVRAPPNRTIFDSDLMSARLGSGPCPPSSNSGISAT 182
Db 158 PPHYQGCTLQLRDPQQLELNRESVRAPPNRTIFDSDLMSARLGSGPCPPSSNSGISAT 217

Qy 183 CYGSGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSK 242
Db 218 CYGSGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSK 277

Qy 243 EKDKQKGHPL 252
Db 278 EKDKQKGHPL 287

RESULT 11
US-10-295-027-128
; Sequence 128, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 128
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-295-027-128

Query Match          99.2%; Score 250; DB 15; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.5e-229;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 ELEFVQIIIIIVMMVMVVITCLLSHYKLSARSFISRHSGQRRREDALSSGCLWPSES 62
Db 38 ELEFVQIIIIIVMMVMVVITCLLSHYKLSARSFISRHSGQRRREDALSSGCLWPSES 97
```



Db 38 ELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPSE 97  
QY 63 TVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLOHEIDLPTISLSDGE 122  
Db 98 TVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLOHEIDLPTISLSDGE 157  
QY 123 PPPYQGPCTQLRDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGISAT 182  
Db 158 PPPYQGPCTQLRDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGISAT 217  
QY 183 CYGSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSILLEGTRLHHTHIAPLESAAIWSK 242  
Db 218 CYGSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSILLEGTRLHHTHIAPLESAAIWSK 277  
QY 243 EKDKQKGHPL 252  
Db 278 EKDKQKGHPL 287

RESULT 12

US-10-390-045-11  
; Sequence 11, Application US/10390045  
; Publication No. US20030170713A1  
; GENERAL INFORMATION:  
; APPLICANT: SRIVASTAVA, SHIV  
; APPLICANT: MOUL, JUDD W.  
; APPLICANT: XU, LINDA L.  
; APPLICANT: SEGAWA, TAKEHIKO  
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED  
; FILE REFERENCE: POYNUCLEOTIDE ARRAY  
; CURRENT APPLICATION NUMBER: US/10/390,045  
; CURRENT FILING DATE: 2003-03-18  
; PRIOR APPLICATION NUMBER: US/09/769,482  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,772  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/179,045  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 249  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-390-045-11

Query Match 98.8%; Score 249; DB 14; Length 249;  
Best Local Similarity 100.0%; Pred. No. 1.2e-228;  
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPSE 61  
Db 1 AELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPSE 60  
QY 62 STVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLOHEIDLPTISLSDGE 121  
Db 61 STVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLOHEIDLPTISLSDGE 120  
QY 122 EPPPYQGPCTQLRDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGISAT 181  
Db 121 EPPPYQGPCTQLRDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGISAT 180  
QY 182 TCYSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSILLEGTRLHHTHIAPLESAAIWS 241  
Db 181 TCYSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSILLEGTRLHHTHIAPLESAAIWS 240  
QY 242 EKDKQKGH 250  
Db 241 EKDKQKGH 249

RESULT 13

US-10-434-479-11  
; Sequence 11, Application US/10434479  
; Publication No. US20040092469A1  
; GENERAL INFORMATION:  
; APPLICANT: SRIVASTAVA, SHIV  
; APPLICANT: MOUL, JUDD W.  
; APPLICANT: XU, LINDA L.  
; TITLE OF INVENTION: ANDROGEN-REGULATED PMEPAL GENE AND POLYPEPTIDES  
; FILE REFERENCE: 04995.0057-02000  
; CURRENT APPLICATION NUMBER: US/10/434,479  
; CURRENT FILING DATE: 2003-05-09  
; PRIOR APPLICATION NUMBER: 10/390,045  
; PRIOR FILING DATE: 2003-03-18  
; PRIOR APPLICATION NUMBER: 09/769,482  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,772  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/179,045  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 249  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-434-479-11

Query Match 98.8%; Score 249; DB 16; Length 249;  
Best Local Similarity 100.0%; Pred. No. 1.2e-228;  
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPSE 61  
Db 1 AELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPSE 60  
QY 62 STVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLOHEIDLPTISLSDGE 121  
Db 61 STVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLOHEIDLPTISLSDGE 120  
QY 122 EPPPYQGPCTQLRDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGISAT 181  
Db 121 EPPPYQGPCTQLRDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGISAT 180  
QY 182 TCYSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSILLEGTRLHHTHIAPLESAAIWS 241  
Db 181 TCYSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSILLEGTRLHHTHIAPLESAAIWS 240  
QY 242 EKDKQKGH 250  
Db 241 EKDKQKGH 249

RESULT 14

US-10-000-256A-158  
; Sequence 158, Application US/10000256A  
; Publication No. US20030039983A1  
; GENERAL INFORMATION:  
; APPLICANT: Sun, Yongming  
; APPLICANT: Recipon, Herve  
; APPLICANT: Chen, Sei-Yu  
; APPLICANT: Liu, Chenghua  
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes ar  
; FILE REFERENCE: DEX-0259  
; CURRENT APPLICATION NUMBER: US/10/000,256A  
; CURRENT FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: 60/244,782  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 240  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 158  
; LENGTH: 241  
; TYPE: PRT  
; ORGANISM: Homo sapien

US-10-000-256A-158

Query Match 71.8%; Score 181; DB 14; Length 241;  
Best Local Similarity 100.0%; Pred. No. 6.5e-164;  
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 72 PQVAPPRTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDGEEPPPYQGPGCT 131  
Db 60 PQVAPPRTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDGEEPPPYQGPGCT 119  
Qy 132 LQLRDPEQQLNRESVRAPPNRTIFDSDLMDSARLGPCPPSSNSGISATCYGSGGRME 191  
Db 120 LQLRDPEQQLNRESVRAPPNRTIFDSDLMDSARLGPCPPSSNSGISATCYGSGGRME 179  
Qy 192 GPPPTYSEVIGHYPGSSFQHQSSGPPSLLGTRLHHTHIAPLESAAIWSKDKQKGHP 251  
Db 180 GPPPTYSEVIGHYPGSSFQHQSSGPPSLLGTRLHHTHIAPLESAAIWSKDKQKGHP 239  
Qy 252 L 252  
Db 240 L 240

RESULT 15

US-09-796-753-58  
; Sequence 58, Application US/09796753  
; Publication No. US20030027998A1  
; GENERAL INFORMATION:  
; APPLICANT: McCarthy, Sean A.  
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF  
; FILE REFERENCE: 7853-227-999  
; CURRENT APPLICATION NUMBER: US/09/796,753  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 09/183,175  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: 09/223,094  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/223,546  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/224,246  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/259,388  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR APPLICATION NUMBER: 60/122,458  
; PRIOR FILING DATE: 1999-03-01  
; PRIOR APPLICATION NUMBER: 09/312,359  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 09/336,536  
; PRIOR FILING DATE: 1999-06-18  
; PRIOR APPLICATION NUMBER: 09/342,687  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: 09/345,464  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: 09/365,164  
; PRIOR FILING DATE: 1999-07-30  
; PRIOR APPLICATION NUMBER: 09/399,723  
; PRIOR FILING DATE: 1999-09-20  
; PRIOR APPLICATION NUMBER: 09/409,634  
; PRIOR FILING DATE: 1999-09-30  
; PRIOR APPLICATION NUMBER: 09/471,179  
; PRIOR FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: 09/474,071  
; PRIOR FILING DATE: 1999-12-29  
; PRIOR APPLICATION NUMBER: 09/474,072  
; PRIOR FILING DATE: 1999-12-29  
; PRIOR APPLICATION NUMBER: 09/514,010  
; PRIOR FILING DATE: 2000-02-25  
; PRIOR APPLICATION NUMBER: 09/516,745  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 09/572,002  
; PRIOR FILING DATE: 2000-05-14  
; PRIOR APPLICATION NUMBER: 09/597,993  
; PRIOR FILING DATE: 2000-06-19

; PRIOR APPLICATION NUMBER: 09/599,596  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 09/630,334  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: 09/606,565  
; PRIOR FILING DATE: 2000-06-29  
; PRIOR APPLICATION NUMBER: 09/606,317  
; PRIOR FILING DATE: 2000-06-29  
; PRIOR APPLICATION NUMBER: 09/665,666  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: 09/677,751  
; PRIOR FILING DATE: 2000-09-30  
; NUMBER OF SEQ ID NOS: 162  
; SEQ ID NO 58  
; LENGTH: 217  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-796-753-58

Query Match 20.2%; Score 51; DB 10; Length 217;  
Best Local Similarity 100.0%; Pred. No. 3.6e-40;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 111 LPPTISLSDGEEPPPYQGPGCTLQLRDPEQQLNRESVRAPPNRTIFDSDL 161  
Db 81 LPPTISLSDGEEPPPYQGPGCTLQLRDPEQQLNRESVRAPPNRTIFDSDL 131

Search completed: May 20, 2004, 15:42:13  
Job time : 49 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2004, 15:32:40 ; Search time 21 Seconds  
(without alignments)  
1154.298 Million cell updates/sec

Title: US-09-857-826B-17  
Perfect score: 252  
Sequence: 1 MALEFVQIIIIIVVMVMV.....PLESAAIWSKEKDKQKGHPL 252

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR\_78: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.2	308	2 A4496	transcription repr
2	8	3.2	582	2 S33937	hexon-associated p
3	8	3.2	794	2 T36972	probable membrane
4	8	3.2	1279	1 E69681	peptide synthetase
5	7	2.8	102	2 S04126	probable phospholi
6	7	2.8	110	4 S49808	hypothetical prote
7	7	2.8	116	2 A72685	hypothetical prote
8	7	2.8	161	2 E81245	conserved hypothet
9	7	2.8	220	2 A25513	25K silk protein -
10	7	2.8	267	2 T20096	hypothetical prote
11	7	2.8	271	2 JQ0950	ICP 18.5 protein -
12	7	2.8	279	2 G86406	unknown protein (i
13	7	2.8	289	2 I60170	syntaxin 3A - mous
14	7	2.8	296	2 T48926	bZIP transcription
15	7	2.8	297	2 AF2980	conserved hypothet
16	7	2.8	297	2 H98302	hypothetical prote
17	7	2.8	334	2 B87545	hypothetical prote
18	7	2.8	339	2 S24161	lipase B - Pseudom
19	7	2.8	375	2 A75264	hypothetical prote
20	7	2.8	387	2 H72299	hypothetical prote
21	7	2.8	410	2 C96803	hypothetical prote
22	7	2.8	447	2 T16527	hypothetical prote
23	7	2.8	450	2 A57426	puc expression act
24	7	2.8	558	2 AE3580	iron(III)-transpor
25	7	2.8	596	2 JC4697	trehalose trehaloh
26	7	2.8	603	2 T37518	probable transcrip
27	7	2.8	620	2 G90488	hypothetical prote
28	7	2.8	659	2 E84176	DNA mismatch repai
29	7	2.8	685	2 T09159	RNA helicase prh75

30	7	2.8	721	2 T45495	probable transposa
31	7	2.8	721	2 A82941	hypothetical prote
32	7	2.8	778	2 T30430	hypothetical prote
33	7	2.8	898	2 T05099	hypothetical prote
34	7	2.8	952	2 T18837	hypothetical prote
35	7	2.8	975	2 T16073	hypothetical prote
36	7	2.8	2139	2 T18296	myosin heavy chain
37	7	2.8	2500	1 WMH082	HIV-BP2 enhancer-b
38	7	2.8	26926	1 I38344	titin, cardiac mus
39	6	2.4	20	2 I64036	hypothetical prote
40	6	2.4	47	2 D81117	hypothetical prote
41	6	2.4	61	2 H83614	hypothetical prote
42	6	2.4	66	2 F69161	conserved hypothet
43	6	2.4	71	2 I58381	receptor tyrosine
44	6	2.4	76	2 S75641	hypothetical prote
45	6	2.4	80	2 D75410	conserved hypothet

ALIGNMENTS

RESULT 1

A4496  
transcription repressor factor AEF-1, adult enhancer factor 1 - fruit fly (Drosophila m  
C;Species: Drosophila melanogaster  
C;Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 16-Feb-1997  
C;Accession: A4496  
R;Falb, D.; Maniatis, T.  
Mol. Cell. Biol. 12, 4093-4103, 1992  
A;Title: Drosophila transcriptional repressor protein that binds specifically to negati  
A;Reference number: A4496; MUID:92375078; PMID:1508206  
A;Accession: A4496  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-308 <FAL>  
A;Experimental source: ovary  
A;Note: sequence extracted from NCBI backbone (NCBIN:111713, NCBIP:111714)  
C;Genetics:  
A;Gene: FlyBase:Aef1  
A;Cross-references: FlyBase:FBgn0005694  
C;Keywords: DNA binding; transcription regulation

Query Match 3.2%; Score 8; DB 2; Length 308;  
Best Local Similarity 100.0%; Pred. No. 4.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 212 QQSSGPPS 219

Db 75 QQSSGPPS 82

RESULT 2

S33937

hexon-associated protein - human adenovirus 12  
C;Species: Mastadenovirus h12 (human adenovirus 12)  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999  
C;Accession: S33937  
R;Sprenkel, J.  
Submitted to the EMBL Data Library, June 1993  
A;Reference number: S33928  
A;Accession: S33937  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-582 <SPR>  
A;Cross-references: EMBL:X73487; NID:g313361; PIDN:CAA51886.1; PID:g313371  
C;Superfamily: adenovirus peripentonal hexon-associated protein

Query Match 3.2%; Score 8; DB 2; Length 582;  
Best Local Similarity 100.0%; Pred. No. 8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 90 AQRERFHR 97

|||||



Db 134 AQERFHR 141

RESULT 3

T36972  
probable membrane associated protein - Streptomyces coelicolor (fragment)  
C;Species: Streptomyces coelicolor  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C;Accession: T36972  
R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, August 1999  
A;Reference number: 221618  
A;Accession: T36972  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-794 <OLI>  
A;Cross-references: EMBL:AL109949; PIDN:CAB52886.1; GSPDB:GN00070; SCORDB:SCJ11.01c  
A;Experimental source: strain A3(2)  
C;Genetics:  
A;Gene: SCORDB:SCJ11.01c

Query Match 3.2%; Score 8; DB 2; Length 794;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 142 ELNRESVR 149

Db 646 ELNRESVR 653

RESULT 4

E69681  
peptide synthetase ppsB - Bacillus subtilis  
C;Species: Bacillus subtilis  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 03-Nov-2000  
C;Accession: E69681  
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd  
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Serot  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A;Reference number: A69580; MUID:98044033; PMID:9384377  
A;Accession: E69681  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-1279 <KUN>  
A;Cross-references: GB:Z99113; GB:AL009126; NID:g2634090; PIDN:CAB13713.1; PID:g2634213  
A;Experimental source: strain 168  
C;Genetics:  
A;Gene: ppsB  
C;Superfamily: peptide synthetase ppsB; acetate-CoA ligase homology; acyl carrier protei  
C;Keywords: carrier protein; phosphopantetheine; phosphoprotein  
F;522-960/Domain: acetate-CoA ligase homology <ACL>  
F;978-1046/Domain: acyl carrier protein homology <ACP>  
F;1010/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 3.2%; Score 8; DB 1; Length 1279;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 242 KEKDQKG 249

Db 118 KEKDQKG 125

RESULT 5

S04126  
probable phospholipid transfer protein precursor, aleurone-specific - barley  
N;Alternate names: 10K protein  
C;Species: Hordeum vulgare (barley)  
C;Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Sep-1997  
C;Accession: S04126; S14610; S31457  
R;Jakobsen, K.; Klemsdal, S.S.; Aalen, R.B.; Bosnes, M.; Alexander, D.; Olsen, O.A.  
Plant Mol. Biol. 12, 285-293, 1989  
A;Title: Barley aleurone cell development: molecular cloning of aleurone-specific cDNA  
A;Reference number: S04126  
A;Accession: S04126  
A;Molecule type: mRNA  
A;Residues: 1-102 <JAK>  
A;Cross-references: EMBL:X15257; NID:g18892; PID:g18893  
R;Kalla, R.; Klemsdal, S.S.; Linnestad, C.; Loenneborg, A.; Olsen, O.A.  
submitted to the EMBL Data Library, January 1991  
A;Description: A novel phospholipid transfer protein from immature barley aleurone cel  
A;Reference number: S14610  
A;Accession: S14610  
A;Molecule type: mRNA  
A;Residues: 1-102 <KAL>  
A;Cross-references: EMBL:X57270; NID:g19042; PID:g19043  
R;Kaller, R.; Potter, R.; Nielsen, P.S.; Linnestad, C.; Gabrielsen, O.S.; Olsen, O.A.  
submitted to the EMBL Data Library, December 1992  
A;Description: The promoter of the barley Ltp2 gene encoding an aleurone specific 7kD  
A;Reference number: S31457  
A;Accession: S31457  
A;Molecule type: DNA  
A;Residues: 1-102 <KAW>  
A;Cross-references: EMBL:X69793; NID:g683753; PID:g19041  
F;1-28/Domain: (or 31) signal sequence #status predicted <SIG>  
F;29-102/Product: (or 32-102) probable phospholipid transfer protein, aleurone-specifi.

Query Match 2.8%; Score 7; DB 2; Length 102;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VMVMVMV 21

Db 17 VMVMVMV 23

RESULT 6

S49808  
hypothetical protein YM9958.10 - yeast (Saccharomyces cerevisiae)  
C;Species: Saccharomyces cerevisiae  
C;Date: 13-Jan-1995 #sequence\_revision 05-Sep-1996 #text\_change 20-Oct-2000  
C;Accession: S49808  
R;Devlin, K.; Churcher, C.  
submitted to the EMBL Data Library, November 1994  
A;Reference number: S49800  
A;Accession: S49808  
A;Status: conceptual translation of pseudogene  
A;Molecule type: DNA  
A;Residues: 1-110 <DEV>  
A;Cross-references: EMBL:Z46729; NID:g577134; PIDN:CAA86723.1; PID:g577144  
C;Comment: There is no evidence that this sequence is expressed.  
C;Genetics:  
A;Map position: 13  
C;Keywords: pseudogene

Query Match 2.8%; Score 7; DB 4; Length 110;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 IIIIVVV 15

Db 19 IIIIVVV 25

## RESULT 7

A72685  
hypothetical protein APE0897 - Aeropyrum pernix (strain KI)  
C;Species: Aeropyrum pernix  
C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jun-2000  
C;Accession: A72685  
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; KDNA Res. 6, 83-101, 1999  
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum  
A;Reference number: A72450; MUID:99310339; PMID:10382966  
A;Accession: A72685  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-116 <KAW>  
A;Cross-references: DDBJ:AP000060; NID:G5104188; PIDN:BAA79881.1; PID:d1043667; PID:G5104188  
A;Experimental source: strain KI  
C;Genetics:  
A;Gene: APE0897  
C;Superfamily: Aeropyrum pernix hypothetical protein APE0897

Query Match 2.8%; Score 7; DB 2; Length 116;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 SARLGGP 170  
Db 50 SARLGGP 56

## RESULT 8

E81245  
conserved hypothetical protein NMB0043 [imported] - Neisseria meningitidis (strain MC58)  
C;Species: Neisseria meningitidis  
C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Feb-2001  
C;Accession: E81245; D82024  
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M. Science 287, 1809-1815, 2000  
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Venzon  
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
A;Reference number: A81000; MUID:20175755; PMID:10710307

A;Accession: E81245  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-161 <TET>  
A;Cross-references: GB:AE002364; GB:AE002098; NID:G7225269; PIDN:AAF40514.1; PID:G7225269  
A;Experimental source: serogroup B, strain MC58  
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, Nature 404, 502-506, 2000  
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
A;Reference number: A81775; MUID:20222556; PMID:10761919  
A;Accession: D82024  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-161 <PAR>  
A;Cross-references: GB:AL162752; GB:AL157959; NID:G7378778; PIDN:CAB83596.1; PID:G737905  
A;Experimental source: serogroup A, strain Z2491  
C;Genetics:  
A;Gene: NMB0043; NMA0289  
C;Superfamily: Aquifex aeolicus conserved hypothetical protein aq\_1996

Query Match 2.8%; Score 7; DB 2; Length 161;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 144 NRESVRA 150  
Db 134 NRESVRA 140

## RESULT 9

A25513  
25K silk protein - silkworm  
C;Species: Bombyx mori (silkworm)  
C;Date: 30-Jun-1988 #sequence\_revision 31-Dec-1988 #text\_change 21-Jul-2000  
C;Accession: A25513; A23020  
R;Chevillard, M.; Couble, P.; Prudhomme, J.C.  
Nucleic Acids Res. 14, 6341-6342, 1986  
A;Title: Complete nucleotide sequence of the gene encoding the Bombyx mori silk protei  
A;Reference number: A25513; MUID:86312902; PMID:3748814  
A;Accession: A25513  
A;Molecule type: DNA  
A;Residues: 1-220 <CHE>  
A;Cross-references: GB:X04226; GB:X02223; NID:G5898; PIDN:CAA27804.1; PID:G5899  
R;Couble, P.; Chevillard, M.; Moine, A.; Ravel-Chapuis, P.; Prudhomme, J.C.  
Nucleic Acids Res. 13, 1801-1814, 1985  
A;Title: Structural organization of the P-25- gene of Bombyx mori and comparative anal  
A;Reference number: A23020; MUID:85215581; PMID:2987837  
A;Accession: A23020  
A;Molecule type: protein  
A;Residues: 1-41, 'I', 43-81, 'Q', 83 <COU>  
C;Genetics:  
A;Introns: 18/1; 91/2; 137/1; 175/1  
C;Superfamily: silkworm 25K silk protein

Query Match 2.8%; Score 7; DB 2; Length 220;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 ARSFISR 40  
Db 205 ARSFISR 211

## RESULT 10

T20096  
hypothetical protein C50B8.5 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
C;Accession: T20096  
R;Percy, C.  
submitted to the EMBL Data Library, July 1996  
A;Reference number: Z19223  
A;Accession: T20096  
A;Status: preliminary; translated from GB/EMBL/DBSJ  
A;Molecule type: DNA  
A;Residues: 1-267 <WIL>  
A;Cross-references: EMBL:Z77654; PIDN:CAB01128.1; GSPDB:GN00023; CESP:C50B8.5  
A;Experimental source: clone C50B8  
C;Genetics:  
A;Gene: CESP:C50B8.5  
A;Map position: 5  
A;Introns: 23/3; 118/3; 157/3  
C;Superfamily: Caenorhabditis elegans hypothetical protein C50B8.5

Query Match 2.8%; Score 7; DB 2; Length 267;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 193 PPPTYSE 199  
Db 234 PPPTYSE 240

## RESULT 11

JQ0950  
ICP 18.5 protein - infectious laryngotracheitis virus (fragment)  
C;Species: infectious laryngotracheitis virus  
A;Note: host Gallus gallus (chicken)  
C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 03-Nov-2000  
C;Accession: JQ0950  
R;Griffin, A.M.  
submitted to JIPID, January 1991

A;Reference number: JQ0950  
A;Accession: JQ0950  
A;Molecule type: DNA  
A;Residues: 1-271 <GRI>  
A;Experimental source: strain thorne V882  
C;Genetics:

A;Gene: icp18.5  
C;Superfamily: herpesvirus infected cell protein ICP18.5

Query Match 2.8%; Score 7; DB 2; Length 271;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 KLSARSP 37  
| | | | |  
Db 1 KLSARSP 7

# RESULT 12

G86406  
unknown protein [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C;Accession: G86406  
R;Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: G86406  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-279 <STO>  
A;Cross-references: GB:AE005172; NID:g11024879; PIDN:AAG26963.1; GSPDB:GN00141  
C;Genetics:  
A;Map position: 1

Query Match 2.8%; Score 7; DB 2; Length 279;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 VITCLLS 28  
| | | | |  
Db 5 VITCLLS 11

# RESULT 13

I60170  
syntaxin 3A - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 24-Sep-1999  
C;Accession: I60170  
R;Ibaraki, K.; Horikawa, H.P.; Morita, T.; Mori, H.; Sakimura, K.; Mishina, M.; Saisu, H.  
Biochem. Biophys. Res. Commun. 211, 997-1005, 1995  
A;Title: Identification of four different forms of syntaxin 3.  
A;Reference number: I60170; MUID:95321968; PMID:7598732  
A;Accession: I60170  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-289 <RES>  
A;Cross-references: GB:D29797; NID:g924267; PIDN:BAA06180.1; PID:g924268  
C;Superfamily: syntaxin

Query Match 2.8%; Score 7; DB 2; Length 289;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 IIIIVVW 15  
| | | | |  
Db 265 IIIIVVW 271

# RESULT 14

T48926  
bZIP transcription factor-like protein - Arabidopsis thaliana  
N;Alternate names: protein F14L2.10  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
C;Accession: T48926  
R;Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Rudd, S.  
submitted to the Protein Sequence Database, April 2000  
A;Reference number: Z25008  
A;Accession: T48926  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-296 <JOR>  
A;Cross-references: EMBL:AL353818; GSPDB:GN00061; ATSP:F14L2.10  
A;Experimental source: cultivar Columbia; BAC clone F14L2  
C;Genetics:  
A;Gene: ATSP:F14L2.10  
A;Map position: 3  
A;Introns: 269/3

Query Match 2.8%; Score 7; DB 2; Length 296;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 PPSSNSG 178  
| | | | |  
Db 125 PPSSNSG 131

# RESULT 15

AF2980  
conserved hypothetical protein Atu3447 [imported] - Agrobacterium tumefaciens (strain  
C;Species: Agrobacterium tumefaciens  
C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
C;Accession: AF2980  
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCl  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam  
ster, B.W.  
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A;Reference number: AB2577; MUID:21608550; PMID:11743193  
A;Accession: AF2980  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-297 <KUR>  
A;Cross-references: GB:AE008689; PIDN:AAL44260.1; PID:g17741845; GSPDB:GN00187  
A;Experimental source: strain C58 (Dupont)  
C;Genetics:  
A;Gene: Atu3447  
A;Map position: linear chromosome

Query Match 2.8%; Score 7; DB 2; Length 297;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 RRREDAL 51  
| | | | |  
Db 34 RRREDAL 40

Search completed: May 20, 2004, 15:36:49  
Job time : 24 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2004, 15:25:50 ; Search time 18 Seconds  
(without alignments)  
728.982 Million cell updates/sec

Title: US-09-857-826B-17  
Perfect score: 252  
Sequence: 1 MALEFVQIIIIIVVMVMV.....PLESAAIWSKEKDQKGHPL 252

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	250	99.2	287	1	TMEP HUMAN
2	51	20.2	260	1	TMEP MOUSE
3	35	13.9	306	1	C181_HUMAN
4	8	3.2	308	1	AEP1_DROME
5	8	3.2	582	1	HEX3_ADE12
6	7	2.8	102	1	NLT2_HORVU
7	7	2.8	145	1	CRAA_ORYLA
8	7	2.8	220	1	SI25_BOMMO
9	7	2.8	289	1	STX3_MOUSE
10	7	2.8	596	1	TRE2_RHISP
11	7	2.8	603	1	YAO7_SCHPO
12	7	2.8	659	1	MUTL_HALN1
13	7	2.8	708	1	TRAB_HUMAN
14	7	2.8	778	1	GSH1_CLOPE
15	7	2.8	925	1	NPA3_MOUSE
16	7	2.8	933	1	NPA3_HUMAN
17	7	2.8	952	1	YK15_CAEEL
18	7	2.8	1833	1	ZEP2_HUMAN
19	6	2.4	79	1	YF71_HAEIN
20	6	2.4	86	1	YBBD_ECOLI
21	6	2.4	90	1	IM10_HUMAN
22	6	2.4	95	1	GON1_PAGMA
23	6	2.4	103	1	RL44_AEDTR
24	6	2.4	119	1	ACL4_ACHLY
25	6	2.4	119	1	MCD1_HUMAN
26	6	2.4	128	1	Y962_METJA
27	6	2.4	131	1	SOH1_HUMAN
28	6	2.4	131	1	SOH1_MOUSE
29	6	2.4	139	1	SOH1_SCHPO
30	6	2.4	140	1	ZG49_XENLA
31	6	2.4	141	1	X_WHV1
32	6	2.4	141	1	X_WHV59
33	6	2.4	141	1	X_WHVW6

34	6	2.4	142	1	IBPB_ECOLI
35	6	2.4	142	1	MRAZ_CLOAB
36	6	2.4	144	1	NDK_LYCES
37	6	2.4	148	1	SOH1_TABSO
38	6	2.4	148	1	YPX2_BLVJ
39	6	2.4	149	1	Y4OM_RHISN
40	6	2.4	150	1	YF87_MYCPN
41	6	2.4	151	1	TSR3_HUMAN
42	6	2.4	151	1	Y273_METTH
43	6	2.4	155	1	MSRB_XYLPA
44	6	2.4	155	1	MSRB_XYLFT
45	6	2.4	163	1	DEF3_SHEON

P29210	escherichia
Q97h80	clostridium
P47921	lycopersico
Q8mpd6	taenia soli
P03413	bovine leuk
P55598	rhizobium s
P75193	mycoplasma
Q8izw5	homo sapien
Q26373	methanobact
Q9pf29	xylella fas
Q87aj9	xylella fas
Q8ee60	shewanella

ALIGNMENTS

RESULT 1  
TMEP\_HUMAN  
ID TMEP\_HUMAN STANDARD; PRT; 287 AA.  
AC Q969W9; Q96B72; Q9UJD3;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Transmembrane prostate androgen-induced protein (Solid tumor-associated protein).  
GN TMEPAL OR PMEPA1 OR STAG1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RX MEDLINE=20334621; PubMed=10873380;  
RA Xu L.L., Shanmugam N., Segawa T., Sesterhenn I.A., McLeod D.G., Moul J.W., Srivastava S.;  
RT "A novel androgen-regulated gene, PMEPA1, located on chromosome 20q13 exhibits high level expression in prostate.";  
RL Genomics 66:257-263 (2000).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=21453682; PubMed=11568975;  
RA Rae F.K., Hooper J.D., Nicol D.L., Clements J.A.;  
RT "Characterization of a novel gene, STAG1/PMEPA1, upregulated in renal cell carcinoma and other solid tumors.";  
RL Mol. Carcinog. 32:44-53 (2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L., Bailey J., Barlow K.P., Bates K.N., Beard L.M., Beare D.M., Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J., Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Clegg S., Clegg S., Clegg S., Collier R.E., Connor R.E., Corby N.R., Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Leharveslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D., Marsh V.L., Martin S.L., McConachie L.J., McLeay K., McMurray A.A., Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I., Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H., Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S., Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E., Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A., Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,

RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20.";  
RL Nature 414:865-871(2001).  
RN [4]  
RP SEQUENCE OF 8-287 FROM N.A. (ISOFORM 1).  
RC TISSUE=Kidney;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- SUBUNIT: Interacts with the WW domains of NEDD4 (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type Ib membrane protein (Potential).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q969W9-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q969W9-2; Sequence=VSP\_006438;  
CC Note=No experimental confirmation available;  
CC TISSUE SPECIFICITY: Highest expression in prostate. Also expressed  
in ovary.  
CC -!- INDUCTION: By androgen.  
CC -!- SIMILARITY: BELONGS TO THE TMEM165 FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL; AF224278; AAF86322.1; -  
DR EMBL; AF305616; AAL16781.1; -  
DR EMBL; AF305426; AAL09357.1; -  
DR EMBL; AL035541; CAB55862.1; -  
DR EMBL; BC015918; AAL15918.1; ALT\_INIT.  
DR Genew; HGNC:14107; TMEM165.  
DR MIM; 606564; -  
DR GO; GO:0016021; C:integral to membrane; NAS.  
DR GO; GO:0030521; P:androgen receptor signaling pathway; NAS.  
KW Transmembrane; Alternative splicing.  
FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 41 63 POTENTIAL.  
FT DOMAIN 64 287 CYTOPLASMIC (POTENTIAL).  
FT SITE 158 161 WW-BINDING (POTENTIAL).  
FT SITE 229 232 WW-BINDING (POTENTIAL).  
FT VARSPLIC 1 37 MRLMGVNSTAAAGQPNVSTCNCKRSLFQSMET ->  
FT MA (in isoform 2).  
FT /FTId=VSP\_006438.  
SQ SEQUENCE 287 AA; 31609 MW; 6103473561AE08DA CRC64;  
Query Match 99.2%; Score 250; DB 1; Length 287;  
Best Local Similarity 100.0%; Pred. No. 1.6e-254;  
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 BLEFVQIIIIIVMMVMVVVITCLLSHYKLSARSFISRHSGRRRDLSSSEGCLMPSES 62  
DB |||||  
38 BLEFVQIIIIIVMMVMVVVITCLLSHYKLSARSFISRHSGRRRDLSSSEGCLMPSES 97  
QY 63 TVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTPYLQHEIDLPTISLDGEE 122  
DB |||||  
98 TVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTPYLQHEIDLPTISLDGEE 157  
QY 123 PPVYQGPCTQLRDPQQLRLNRESVRAPNRTIFDSDLMDSARLGCGPPSSNSGISAT 182  
DB |||||  
158 PPVYQGPCTQLRDPQQLRLNRESVRAPNRTIFDSDLMDSARLGCGPPSSNSGISAT 217  
QY 183 CYSGGRMEGPPPTTYSEVIGHYPCSSFGHQSSGPPSLLEGTRLHHTHIAPLESAIWSK 242  
DB |||||  
218 CYSGGRMEGPPPTTYSEVIGHYPCSSFGHQSSGPPSLLEGTRLHHTHIAPLESAIWSK 277  
QY 243 EKDKQKGHPL 252  
DB |||||  
278 EKDKQKGHPL 287  
RESULT 2  
TMEM165  
ID TMEM165 STANDARD; PRT; 260 AA.  
AC Q9D7R2; Q9BQ99;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Transmembrane prostate androgen-induced protein (Nedd4 WW domain-  
binding protein 4).  
DE TMEM165 OR N4WBP4.  
GN Mus musculus (Mouse).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Stomach;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [2]  
RP SEQUENCE OF 59-260 FROM N.A.  
RX MEDLINE=20498735; PubMed=11042109;  
RA Jolliffe C.N., Harvey K.F., Haines B.P., Parasivam G., Kumar S.;  
RT "Identification of multiple proteins expressed in murine embryos as  
binding partners for the WW domains of the ubiquitin-protein ligase  
Nedd4.";  
RT Biochem. J. 351:557-565(2000).  
RL -!- SUBUNIT: Interacts with the WW domains of NEDD4.  
CC -!- SUBCELLULAR LOCATION: Type Ib membrane protein (Potential).  
CC -!- SIMILARITY: BELONGS TO THE TMEM165 FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AK008976; BAB26001.1; --  
DR EMBL; AF220208; AAG44247.1; --  
DR MGD; MGI:1929600; N4wbp4.  
DR GO; GO:0016021; C: integral to membrane; ISS.  
DR GO; GO:0005515; F: protein binding; IPI.  
DR GO; GO:0030521; P: androgen receptor signaling pathway; ISS.  
KW Transmembrane.  
FT DOMAIN 1 20 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 21 43 POTENTIAL.  
FT DOMAIN 44 260 CYTOPLASMIC (POTENTIAL).  
FT SITE 136 139 WW-BINDING (POTENTIAL).  
FT SITE 207 210 WW-BINDING (POTENTIAL).  
SQ SEQUENCE 260 AA; 28715 MW; 31AD07BD16B0D77D CRC64;

Query Match 20.2%; Score 51; DB 1; Length 260;  
Best Local Similarity 100.0%; Pred. No. 1.2e-45;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 111 LPPTISLSDGEEPPPYQGPTQLRDPEQQLNRESVRAPPNRTIFDSDL 161  
Db 124 LPPTISLSDGEEPPPYQGPTQLRDPEQQLNRESVRAPPNRTIFDSDL 174

RESULT 3

C181 HUMAN  
ID C181 HUMAN STANDARD; PRT; 306 AA.  
AC O15165; O15166; O15167; O15168;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Protein C18orf1.  
GN C18ORF1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
RC TISSUE=Brain;  
RX MEDLINE=98140124; PubMed=9479497;  
RA Yoshikawa T., Sanders A.R., Esterling L.E., Detera-Wadleigh S.D.;  
RT "Multiple transcriptional variants and RNA editing in C18orf1, a novel  
RT gene with LDLRA and transmembrane domains on 18p11.2";  
RL Genomics 47:246-257(1998).  
CC -!- SUBCELLULAR LOCATION: Type Ib membrane protein (Potential).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=4;  
CC Name=Alpha-1;  
CC IsoId=O15165-1; Sequence=Displayed;  
CC Name=Alpha-2;  
CC IsoId=O15165-2; Sequence=VSP\_006440;  
CC Name=Beta-1;  
CC IsoId=O15165-3; Sequence=VSP\_006439;  
CC Name=Beta-2;  
CC IsoId=O15165-4; Sequence=VSP\_006439, VSP\_006440;  
CC -!- SIMILARITY: BELONGS TO THE TMRPAI FAMILY.  
CC -!- SIMILARITY: Contains 1 LDL-receptor class A domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF009424; AAC52023.1; --

DR EMBL; AF009425; AAC52024.1; --  
DR EMBL; AF009426; AAC52025.1; --  
DR EMBL; AF009427; AAC52026.1; --  
DR HSSP; P01130; IAJJ.  
DR Genew; HGNC:1224; C18orf1.  
DR MIM; 606571; --  
DR GO; GO:0016021; C: integral to membrane; NAS.  
DR InterPro; IPR002172; LDL receptor\_A.  
DR Pfam; PF00057; ldl\_recept\_a; 1.  
DR SMART; SM00192; LDLa; 1.  
DR PROSITE; PS01209; LDLRA\_1; FALSE\_NEG.  
DR PROSITE; PS50068; LDLRA\_2; 1.  
KW Transmembrane; Alternative splicing.  
FT DOMAIN 1 64 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 65 85 POTENTIAL.  
FT DOMAIN 86 306 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 16 48 LDL-RECEPTOR CLASS A.  
FT VARSPLIC 1 61 MPEAGFOATNAPTECKFTCTSGKCLYLGSLVCNQNDGDN  
FT SDEENCLLVTEHPPPGIFNS -> MAA (in isoform  
FT Beta-1 and isoform Beta-2).  
FT /FTId=VSP\_006439.  
FT Missing (in isoform Alpha-2 and isoform  
FT Beta-2).  
FT VARSPLIC 113 130  
SQ SEQUENCE 306 AA; 33900 MW; F48EF66E329201BD CRC64;  
Query Match 13.9%; Score 35; DB 1; Length 306;  
Best Local Similarity 100.0%; Pred. No. 8.4e-29;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 106 QHEIDLPTISLSDGEEPPPYQGPTQLRDPEQQ 140  
Db 163 QHEIDLPTISLSDGEEPPPYQGPTQLRDPEQQ 197  
RESULT 4  
AEF1 DROME  
ID AEF1 DROME STANDARD; PRT; 308 AA.  
AC P39413; Q9VP37;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Adult enhancer factor 1 (AEF-1).  
GN AEF1 OR CG5683.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RX MEDLINE=92375078; PubMed=1508206;  
RA Falb D., Maniatis T.;  
RT "Drosophila transcriptional repressor protein that binds specifically  
RT to negative control elements in fat body enhancers.";  
RL Mol. Cell. Biol. 12:4093-4103(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.P., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,



RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley; TISSUE=Embryo;  
RX MEDLINE=22426066; PubMed=12537569;  
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,  
RA George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,  
RA Rubin G.M., Celniker S.E.;  
RT "A Drosophila full-length cDNA resource.";  
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).  
CC -!- FUNCTION: TRANSCRIPTIONAL REPRESSOR THAT BINDS SPECIFICALLY TO FAT  
CC BODY-SPECIFIC ENHANCERS, NAMELY THE ADULT ADH ENHANCER (AAE) AND  
CC THE ENHANCER THAT CONTROLS YOLK PROTEIN GENE EXPRESSION.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- TISSUE SPECIFICITY: FOUND IN ALL TISSUES EXAMINED INCLUDING THE  
CC OVARY AND THE FAT BODY.  
CC -!- DEVELOPMENTAL STAGE: HIGHER LEVELS ARE FOUND IN THIRD-INSTAR  
CC LARVAE AND IN ADULTS.  
CC -!- SIMILARITY: Contains 4 C2H2-type zinc fingers.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
CC EMBL; M90755; -; NOT ANNOTATED\_CDS.  
DR EMBL; AB003594; AAF51722.1; -;  
DR EMBL; AY060435; AAL25474.1; -;  
DR PIR; A44496; A44496.  
DR HSP; P08046; 1A1H.  
DR TRANSFAC; T01513; -;  
DR FlyBase; FBgn0005694; Aef1.  
DR InterPro; IPR007087; Znf\_C2H2.  
DR Pfam; PF00096; zf-C2H2; 4.  
DR ProDom; PD000003; Znf\_C2H2; 2.  
DR SMART; SM00355; Znf\_C2H2; 4.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 4.  
DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 4.  
KW Transcription regulation; Repressor; DNA-binding; Nuclear protein;  
KW Zinc-finger; Metal-binding; Repeat.  
FT DOMAIN 56 76 GLN-RICH (OPA-REPEAT).  
FT DOMAIN 71 76 POLY-GLN.  
FT DOMAIN 105 127 ALA-RICH.  
FT DOMAIN 132 135 POLY-PRO.  
FT ZN\_FING 184 206 C2H2-TYPE 1.

FT ZN\_FING 212 234 C2H2-TYPE 2.  
FT ZN\_FING 240 262 C2H2-TYPE 3.  
FT ZN\_FING 268 290 C2H2-TYPE 4.  
SQ SEQUENCE 308 AA; 33782 MW; 54F04D90B6FC4250 CRC64;  
  
Query Match 3.2%; Score 8; DB 1; Length 308;  
Best Local Similarity 100.0%; Pred.No. 1.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 212 QQSSGPPS 219  
Db 75 QQSSGPPS 82  
|||||  
  
RESULT 5  
HEX3\_ADE12 STANDARD; PRT; 582 AA.  
AC P36712;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Peripentonal hexon-associated protein (Protein IIIA).  
GN PIIIA.  
OS Human adenovirus type 12.  
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
OX NCBI\_TaxID=28282;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94076430; PubMed=8254750;  
RA Sprengel J., Schmitz B., Heuss-Neitzel D., Zock C., Doerfler W.;  
RT "Nucleotide sequence of human adenovirus type 12 DNA: comparative  
RT functional analysis.";  
RL J. Virol. 68:379-389(1994).  
RN [2]  
RP SEQUENCE OF 566-576.  
RC STRAIN=Huie;  
RX MEDLINE=93174944; PubMed=8438575;  
RA Freimuth P., Anderson C.W.;  
RT "Human adenovirus serotype 12 virion precursors pMu and pVI are  
RT cleaved at amino-terminal and carboxy-terminal sites that conform to  
RT the adenovirus 2 endoproteinase cleavage consensus sequence.";  
RL Virology 193:348-355(1993).  
CC  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
CC EMBL; X73487; CAA51886.1; -;  
DR PIR; S33937; S33937.  
DR InterPro; IPR003479; Hex\_IIIa.  
DR Pfam; PF02455; Hex\_IIIa; 1.  
KW Hexon-associated protein; Late protein.  
SQ SEQUENCE 582 AA; 64786 MW; 8341875AE9A267F1 CRC64;  
  
Query Match 3.2%; Score 8; DB 1; Length 582;  
Best Local Similarity 100.0%; Pred.No. 3.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 90 AQRERFHR 97  
Db 134 AQRERFHR 141  
|||||  
  
RESULT 6  
NLT2\_HORVU STANDARD; PRT; 102 AA.  
ID NLT2\_HORVU  
AC P20145;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)



DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Probable nonspecific lipid-transfer protein precursor (LTP) (Aleuron-  
DE specific 10 kDa protein) (B-FABP).  
GN LTP2 OR B1E OR LTP.  
OS Hordeum vulgare (Barley).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
OC Triticeae; Hordeum.  
OX NCBI\_TaxID=4513;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Bomi;  
RA Jakobsen K., Klemsdal S.S., Aalen R.B., Bosnes M., Alexander D.,  
RA Olsen O.-A.;  
RT "Barley aleurone cell development: molecular cloning of aleurone-  
RT specific cDNAs from immature grains.";  
RL Plant Mol. Biol. 12:285-293(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Bomi;  
RX MEDLINE=95152558; PubMed=7849757;  
RA Kalla R., Shimamoto K., Potter R., Nielsen P., Linnestad C.,  
RA Olsen O.-A.;  
RT "The promoter of the barley aleurone-specific gene encoding a  
RT putative 7 kDa lipid transfer protein confers aleurone cell-specific  
RT expression in transgenic rice.";  
RL Plant J. 6:849-860(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Bomi; TISSUE=Leaf;  
RA Kalla R., Potter R., Nielsen P.S., Linnestad C., Gabrielsen O.S.,  
RA Olsen O.-A.;  
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Potential phospholipid transfer protein.  
CC -!- TISSUE SPECIFICITY: ALEURONE.  
CC -!- DEVELOPMENTAL STAGE: MAXIMUM MRNA ABUNDANCE AROUND MID-PHASE OF  
CC GRAIN DEVELOPMENT.  
CC -!- SIMILARITY: Belongs to the plant LTP family. B1E subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X15257; CAA33329.1; -  
CC EMBL; X57270; CAA40542.1; -  
CC EMBL; X69793; CAA49448.1; -  
CC PIR; S04126; S04126.  
CC InterPro; IPR003612; AAI.  
CC InterPro; IPR000528; Plant\_LTP.  
CC Pfam; PF00234; tryp\_alpha\_aml; 1.  
CC SMART; SM00499; AAI; 1.  
CC PROSITE; PS00597; PLANT\_LTP; FALSE NEG.  
KW Lipid-binding; Transport; Signal; Multigene family.  
FT SIGNAL 1 35  
FT CHAIN 36 102  
FT PROBABLE NONSPECIFIC LIPID-TRANSFER  
FT PROTEIN.  
FT DISULFID 37 71  
FT DISULFID 45 59  
FT DISULFID 60 95  
FT DISULFID 69 102  
FT BY SIMILARITY.  
SQ SEQUENCE 102 AA; 10357 MW; 474EEF35D0DD0558 CRC64;  
Query Match 2.8%; Score 7; DB 1; Length 102;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 15 VMVMVV 21  
Db 17 VMVMVV 23

DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Probable nonspecific lipid-transfer protein precursor (LTP) (Aleuron-  
DE specific 10 kDa protein) (B-FABP).  
GN LTP2 OR B1E OR LTP.  
OS Hordeum vulgare (Barley).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
OC Triticeae; Hordeum.  
OX NCBI\_TaxID=4513;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Bomi;  
RA Jakobsen K., Klemsdal S.S., Aalen R.B., Bosnes M., Alexander D.,  
RA Olsen O.-A.;  
RT "Barley aleurone cell development: molecular cloning of aleurone-  
RT specific cDNAs from immature grains.";  
RL Plant Mol. Biol. 12:285-293(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Bomi;  
RX MEDLINE=95152558; PubMed=7849757;  
RA Kalla R., Shimamoto K., Potter R., Nielsen P., Linnestad C.,  
RA Olsen O.-A.;  
RT "The promoter of the barley aleurone-specific gene encoding a  
RT putative 7 kDa lipid transfer protein confers aleurone cell-specific  
RT expression in transgenic rice.";  
RL Plant J. 6:849-860(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Bomi; TISSUE=Leaf;  
RA Kalla R., Potter R., Nielsen P.S., Linnestad C., Gabrielsen O.S.,  
RA Olsen O.-A.;  
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Potential phospholipid transfer protein.  
CC -!- TISSUE SPECIFICITY: ALEURONE.  
CC -!- DEVELOPMENTAL STAGE: MAXIMUM MRNA ABUNDANCE AROUND MID-PHASE OF  
CC GRAIN DEVELOPMENT.  
CC -!- SIMILARITY: Belongs to the plant LTP family. B1E subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X15257; CAA33329.1; -  
CC EMBL; X57270; CAA40542.1; -  
CC EMBL; X69793; CAA49448.1; -  
CC PIR; S04126; S04126.  
CC InterPro; IPR003612; AAI.  
CC InterPro; IPR000528; Plant\_LTP.  
CC Pfam; PF00234; tryp\_alpha\_aml; 1.  
CC SMART; SM00499; AAI; 1.  
CC PROSITE; PS00597; PLANT\_LTP; FALSE NEG.  
KW Lipid-binding; Transport; Signal; Multigene family.  
FT SIGNAL 1 35  
FT CHAIN 36 102  
FT PROBABLE NONSPECIFIC LIPID-TRANSFER  
FT PROTEIN.  
FT DISULFID 37 71  
FT DISULFID 45 59  
FT DISULFID 60 95  
FT DISULFID 69 102  
FT BY SIMILARITY.  
SQ SEQUENCE 102 AA; 10357 MW; 474EEF35D0DD0558 CRC64;  
Query Match 2.8%; Score 7; DB 1; Length 102;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 15 VMVMVV 21  
Db 17 VMVMVV 23

RESULT 7  
CRAA ORYLA STANDARD; PRT; 145 AA.  
AC O73919;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Alpha crystallin A chain (Fragment).  
GN CRYAA.  
OS Oryzias latipes (Medaka fish) (Japanese ricefish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;  
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.  
OX NCBI\_TaxID=8090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Carolina Biological;  
RA Loosli F., Koester R.W., Carl M., Krone A., Wittbrodt J.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: May contribute to the transparency and refractive index  
CC of the lens.  
CC -!- SIMILARITY: Belongs to the small heat shock protein (HSP20)  
CC family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AJ000940; CAA04397.1; -  
CC InterPro; IPR001436; Crystallin\_alpha.  
CC InterPro; IPR003090; Crystallin\_N.  
CC InterPro; IPR002068; Hsp20.  
CC InterPro; IPR008978; HSP20\_chap.  
CC Pfam; PF00525; crystallin; 1.  
CC Pfam; PF00011; HSP20; 1.  
CC PRINTS; PR00299; ACRYSTALLIN.  
CC ProDom; PD001193; Crystallin\_N; 1.  
CC PROSITE; PS01031; HSP20; 1.  
KW Eye lens protein.  
FT NON\_TER 1 1  
FT NON\_TER 145 145  
SQ SEQUENCE 145 AA; 16480 MW; 7B53C5ADFD19706E CRC64;  
Query Match 2.8%; Score 7; DB 1; Length 145;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 174 SSNSGIS 180  
Db 39 SSNSGIS 45  
RESULT 8  
SI25\_BOMMO STANDARD; PRT; 220 AA.  
ID SI25\_BOMMO  
AC P04148;  
DT 01-NOV-1986 (Rel. 03, Created)  
DT 01-NOV-1986 (Rel. 03, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 25 kDa silk protein precursor (p25).  
GN P25.  
OS Bombyx mori (Silk moth).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;  
OC Bombycidae; Bombyx.  
OX NCBI\_TaxID=7091;



Query Match 2.8%; Score 7; DB 1; Length 289;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 IIIIVVV 15  
 DB 265 IIIIVVV 271

RESULT 10  
 TREZ\_RHISP  
 ID TREZ\_RHISP STANDARD; PRT; 596 AA.  
 AC Q53238;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-  
 DE alpha-D-(1->4)-alpha-D-glucano]trehalose trehalohydrolase)  
 DE (Maltoooligosyl trehalose trehalohydrolase).  
 GN TREZ.  
 OS Rhizobium sp. (strain M-11).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.  
 OX NCBI\_TaxID=391;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96219094; PubMed=8829547;  
 RA Maruta K., Hattori K., Nakada T., Kubota M., Sugimoto T., Kurimoto M.;  
 RT "Cloning and sequencing of trehalose biosynthesis genes from Rhizobium  
 sp. M-11.";  
 RL Biosci. Biotechnol. Biochem. 60:717-720(1996).  
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic  
 CC linkage in 4-alpha-D-(1->4)-alpha-D-glucanosyl}(n) trehalose to  
 CC yield trehalose and alpha-(1->4)-D-glucan.  
 CC -!- PATHWAY: Trehalose biosynthesis.  
 CC -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 EMBL; D78001; BAA11187.1; -.  
 DR PIR; JC4697; JC4697.  
 DR InterPro; IPR006047; Alpha\_amil\_cat.  
 DR Pfam; PF00128; alpha-amylase; 1.  
 DR Hydrolase; Glycosidase.  
 FT ACT\_SITE 265 265 BY SIMILARITY.  
 SQ SEQUENCE 596 AA; 65262 MW; 73BA80AE0534DDCD CRC64;

Query Match 2.8%; Score 7; DB 1; Length 596;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 DALSEGE 55  
 DB 287 DALSEGE 293

RESULT 11  
 YAO7\_SCHPO  
 ID YAO7\_SCHPO STANDARD; PRT; 603 AA.  
 AC Q10086;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Putative transcriptional regulatory protein C11D3.07c.  
 GN SPAC11D3.07C.  
 OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle B.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Foraburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 CC -!- SUBCELLULAR LOCATION: Nuclear (probable).  
 CC -!- SIMILARITY: Contains 1 Zn(2)-Cys(6) fungal-type binuclear cluster  
 CC domain.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 EMBL; Z68166; CAA92308.2; -.  
 DR PIR; T37518; T37518.  
 DR HSSP; P08657; 1CLD.  
 DR GeneDB\_SPombe; SPAC11D3.07c; -.  
 DR InterPro; IPR007219; Fungal trans.  
 DR InterPro; IPR01138; Fungi\_Trn.  
 DR Pfam; PF04082; Fungal trans; 1.  
 DR Pfam; PF00172; Zn clus; 1.  
 DR PRINTS; PR00054; FUNGALZNCYS.  
 DR SMART; SM00066; GAL4; 1.  
 DR PROSITE; PS00463; ZN2\_CY6\_FUNGAL\_1; 1.  
 DR PROSITE; PS50048; ZN2\_CY6\_FUNGAL\_2; 1.  
 KW Hypothetical protein; Transcription regulation; DNA-binding;  
 KW Nuclear protein; Zinc; Metal-binding.  
 FT DNA\_BIND 7 34 ZN(2)-CYS(6), FUNGAL-TYPE (POTENTIAL).  
 SQ SEQUENCE 603 AA; 69815 MW; F4D08CA99B7FF415 CRC64;

Query Match 2.8%; Score 7; DB 1; Length 603;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 PTISLSD 119  
 DB 421 PTISLSD 427

RESULT 12



```
MUTL_HALN1
ID MUTL_HALN1 STANDARD; PRT; 659 AA.
AC Q9HGM6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA mismatch repair protein mutL.
GN MUTL OR VNG0159G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RA "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -!- FUNCTION: This protein is involved in the repair of mismatches in
CC DNA. It is required for dam-dependent methyl-directed DNA mismatch
CC repair. May act as a "molecular matchmaker", a protein that
CC promotes the formation of a stable complex between two or more
CC DNA-binding proteins in an ATP-dependent manner without itself
CC being part of a final effector complex (By similarity).
CC -!- SIMILARITY: Belongs to the DNA mismatch repair mutL/hexB family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB004982; AAG18777.1; -
CC PIR; E84176; E84176.
CC HSSP; P23367; 1BKN.
CC HAMAP; MF 00149; -; 1.
CC InterPro; IPR003594; ATPbind_ATPase.
CC InterPro; IPR002099; DNA_mis_repair.
CC Pfam; PF01119; DNA_mis_repair; 1.
CC Pfam; PF02518; HATPase_c; 1.
CC SMART; SM00387; HATPase_c; 1.
CC TIGRFAMs; TIGR00585; mutL; 1.
CC PROSITE; PS00058; DNA_MISMATCH_REPAIR_1; 1.
KW DNA repair; Complete proteome.
SQ SEQUENCE 659 AA; 67917 MW; 1622B48A4AB5B9B8 CRC64;

Query Match 2.8%; Score 7; DB 1; Length 659;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 GRRREDA 50
DB 651 GRRREDA 657

RESULT 13
TRAB_HUMAN
ID TRAB_HUMAN STANDARD; PRT; 708 AA.
AC Q9UGI0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE TRABID protein (zinc finger RAN-binding domain containing 1).
GN ZRANB1 OR TRABID.

Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Evans P.C., Coadwell W.J., Taylor E., Kilshaw P.J.;
RT "Modulation of TNF alpha signalling by a novel protein TRABID.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Modulates TNF alpha signaling.
CC -!- SIMILARITY: Contains 3 RANBP2-type zinc fingers.
CC -!- SIMILARITY: Contains 1 OTU domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ252060; CAB64449.1; -
CC Genew; HGNC:18224; ZRANB1.
CC InterPro; IPR003323; OTU.
CC InterPro; IPR001876; Znf_RanGDP.
CC Pfam; PF00641; zf-RanBP; 2.
CC SMART; SM00547; Znf_RBZ; 3.
CC PROSITE; PS50802; OTU; 1.
CC PROSITE; PS01358; ZF_RANBP2_1; 3.
CC PROSITE; PS50199; ZF_RANBP2_2; 3.
KW Zinc-finger; Repeat.
FT ZN_FING 3 33 RANBP2-TYPE 1.
FT ZN_FING 84 113 RANBP2-TYPE 2.
FT ZN_FING 149 178 RANBP2-TYPE 3.
FT DOMAIN 432 592 OTU.
SQ SEQUENCE 708 AA; 80906 MW; DE8F7DADD2DF23A CRC64;

Query Match 2.8%; Score 7; DB 1; Length 708;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 116 SLSDGEE 122
DB 695 SLSDGEE 701

RESULT 14
GSH1_CLOPE
ID GSH1_CLOPE STANDARD; PRT; 778 AA.
AC Q8XK30;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable glutamate--cysteine ligase (EC 6.3.2.2) (Gamma-
DE glutamylcysteine synthetase) (Gamma-ECS) (GCS).
GN GSHA OR CPE1573.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX MEDLINE=21664373; PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kubara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
CC -!- CATALYTIC ACTIVITY: ATP + L-glutamate + L-cysteine = ADP +
CC phosphate + gamma-L-glutamyl-L-cysteine.
CC -!- PATHWAY: Glutathione biosynthesis; first step.
CC -!- SIMILARITY: Belongs to the glutamate--cysteine ligase family 1.
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2004, 15:25:51 ; Search time 45 Seconds  
(without alignments)  
1766.902 Million cell updates/sec

Title: US-09-857-826B-17  
Perfect score: 252  
Sequence: 1 MAELBFVQIIIVVMVMV.....PLESAAIWSKEKDKQKGHPL 252

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Listing first 45 summaries

Database : SPTREMBL 25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	237	94.0	237	Q8NER4	Q8ner4 homo sapien
2	53	21.0	53	Q9BYL8	Q9byl8 homo sapien
3	51	20.2	84	Q9NTR9	Q9ntr9 homo sapien
4	35	13.9	200	Q8R5E2	Q8r5e2 mus musculu
5	35	13.9	292	Q8BYE2	Q8bye2 mus musculu
6	35	13.9	306	Q8BWJ4	Q8bwj4 mus musculu
7	8	3.2	490	Q89EU9	Q89eu9 bradyrhizob
8	8	3.2	1117	Q8CK70	Q8ck70 streptomyce
9	8	3.2	1271	P94460	P94460 bacillus su
10	8	3.2	1279	O31827	O31827 bacillus su
11	7	2.8	51	Q8F9I7	Q8f9i7 leptospira
12	7	2.8	61	O97510	O97510 sus scrofa
13	7	2.8	73	Q8VB64	Q8vb64 white spot
14	7	2.8	95	Q86SR9	Q86sr9 homo sapien
15	7	2.8	96	Q8LRN2	Q8lrn2 ginkgo bilo
16	7	2.8	97	Q9PZ10	Q9pz10 xestia c-ni

17	7	2.8	103	10	Q941S3	Q941s3 oryza sativ
18	7	2.8	116	17	Q9YDL6	Q9ydl6 aeropyrum p
19	7	2.8	119	16	Q826M5	Q826m5 streptomyce
20	7	2.8	140	3	Q9P8K2	Q9p8k2 botrytis ci
21	7	2.8	140	11	O35283	O35283 rattus norv
22	7	2.8	145	2	O34222	O34222 vibrio chol
23	7	2.8	152	16	Q8DHA2	Q8dha2 synecococc
24	7	2.8	153	4	Q86US6	Q86us6 homo sapien
25	7	2.8	161	16	Q9JRB6	Q9jrb6 neisseria m
26	7	2.8	173	13	Q9DEV0	Q9dev0 clarias fus
27	7	2.8	176	9	Q9MBK1	Q9mbk1 bacteriopha
28	7	2.8	179	16	Q8PF57	Q8pf57 xanthomonas
29	7	2.8	181	4	Q9NXP9	Q9nxp9 homo sapien
30	7	2.8	182	13	Q7ZUF8	Q7zuf8 brachydanio
31	7	2.8	199	4	Q9Y4L8	Q9y4l8 homo sapien
32	7	2.8	203	17	Q8TQV5	Q8tqv5 methanosarc
33	7	2.8	208	5	Q18725	Q18725 caenorhabdi
34	7	2.8	220	5	Q9BLL8	Q9bll8 bombyx mand
35	7	2.8	227	5	Q9GRX7	Q9grx7 calliphora
36	7	2.8	234	10	Q8S627	Q8s627 oryza sativ
37	7	2.8	239	17	Q976P2	Q976p2 sulfolobus
38	7	2.8	264	16	Q98IE2	Q98ie2 rhizobium l
39	7	2.8	271	12	Q69374	Q69374 gallid herp
40	7	2.8	278	12	Q91RB4	Q91rb4 papaya ring
41	7	2.8	278	12	Q8BD97	Q8bd97 papaya ring
42	7	2.8	278	12	Q8BD78	Q8bd78 papaya ring
43	7	2.8	278	12	Q8BD77	Q8bd77 papaya ring
44	7	2.8	278	12	Q8BD76	Q8bd76 papaya ring
45	7	2.8	278	12	Q8BD40	Q8bd40 papaya ring

ALIGNMENTS

RESULT 1

Q8NER4 ID Q8NER4 PRELIMINARY; PRT; 237 AA.  
AC Q8NER4;  
DT 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
DE PMPAL variant A protein.  
OS Homo sapiens (Human).  
OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Brunschwig E.B., Wilson K., Mack D., Dawson D., Lawrence E.,  
RA Willson J.K.V., Lu S., Nosrati A., Swinler S., Beard L.,  
RA Lutterbaugh J.D., Willis J., Platzer P., Markowitz S.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY128643; AAM89277.1; --  
DR GO; GO:0016021; C:integral to membrane; NAS.  
DR GO; GO:0030521; P:androgen receptor signaling pathway; NAS.  
SQ SEQUENCE 237 AA; 26201 MW; A44A274EAABFD930 CRC64;

Query Match	94.0%	Score 237;	DB 4;	Length 237;
Best Local Similarity	100.0%;	Pred. No. 4e-244;		
Matches 237;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	16	MMVMVVITCLSHYKLSARSFISRHSGQRRRREDALSSEGCLWPSESTVSGNGIPEQVY	75	
Db	1	MMVMVVITCLSHYKLSARSFISRHSGQRRRREDALSSEGCLWPSESTVSGNGIPEQVY	60	
Qy	76	APPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDGEEPPPYQGCTQLR	135	
Db	61	APPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDGEEPPPYQGCTQLR	120	
Qy	136	DPEQQLNRESVRAPPNRTIFDSLMDSARLGGPCPPSSNSGISATCYGSGGRMEGPPPP	195	
Db	121	DPEQQLNRESVRAPPNRTIFDSLMDSARLGGPCPPSSNSGISATCYGSGGRMEGPPPP	180	

QY 196 TYSEVIGHYPGSSFOHQSSGPPSLLEGTRLLHHTHIAPLESAAIWSKDKQKGHPL 252  
Db 181 TYSEVIGHYPGSSFOHQSSGPPSLLEGTRLLHHTHIAPLESAAIWSKDKQKGHPL 237

RESULT 2  
Q9BYL8 PRELIMINARY; PRT; 53 AA.  
AC Q9BYL8;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE DJ1059L7.1.2 (Androgen induced type 1b transmembrane protein (PMEPAL), isoform 2) (Fragment).  
GN TMEPAL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Skuce C.;  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL121913; CAC32857.1; --  
DR GO; GO:0016021; C:integral to membrane; IEA.  
KW Transmembrane.  
FT NON TER 53  
SQ SEQUENCE 53 AA; 6064 MW; A3B563FB81F69782 CRC64;

Query Match 21.0%; Score 53; DB 4; Length 53;  
Best Local Similarity 100.0%; Pred. No. 2.2e-48;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAELEFVQIIIIIVVMVMVVTCLLSHYKLSARSFISRHSGRRREDALSS 53  
Db 1 MAELEFVQIIIIIVVMVMVVTCLLSHYKLSARSFISRHSGRRREDALSS 53

RESULT 3  
Q9NTR9 PRELIMINARY; PRT; 84 AA.  
AC Q9NTR9;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE DJ1059L7.1.1 (Androgen induced type 1b transmembrane protein (PMEPAL), isoform 1) (Fragment).  
GN TMEPAL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Skuce C.;  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL121913; CAB88144.1; --  
DR GO; GO:0016021; C:integral to membrane; IEA.  
KW Transmembrane.  
FT NON TER 84  
SQ SEQUENCE 84 AA; 9236 MW; 7FCFCF1FFBFBFC92 CRC64;

Query Match 20.2%; Score 51; DB 4; Length 84;  
Best Local Similarity 100.0%; Pred. No. 4.6e-46;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ELEFVQIIIIIVVMVMVVTCLLSHYKLSARSFISRHSGRRREDALSS 53  
Db 34 ELEFVQIIIIIVVMVMVVTCLLSHYKLSARSFISRHSGRRREDALSS 84

RESULT 4  
Q9R5E2 PRELIMINARY; PRT; 200 AA.  
AC Q9R5E2;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Similar to chromosome 18 open reading frame 1 (Fragment).  
GN D18ERTD653E OR D330030L18RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC022716; AAH22716.1; --  
DR MGD; MGI:1277150; D18Ertd653e.  
FT NON TER 1  
SQ SEQUENCE 200 AA; 22117 MW; 14A78F6E8A4C09A5 CRC64;

Query Match 13.9%; Score 35; DB 11; Length 200;  
Best Local Similarity 100.0%; Pred. No. 1.2e-28;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 QHEIDLPTISLSDGEEPPPYQGPPCTQLQRDPEQQ 140  
Db 57 QHEIDLPTISLSDGEEPPPYQGPPCTQLQRDPEQQ 91

RESULT 5  
Q8BYE2 PRELIMINARY; PRT; 292 AA.  
AC Q8BYE2;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Protein C18ORF1 homolog.  
GN D18ERTD653E OR D330030L18RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Thymus;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium.  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."  
RL Nature 420:563-573(2002).  
DR EMBL; AK040289; BAC30562.1; --  
DR MGD; MGI:1277150; D18Ertd653e.  
SQ SEQUENCE 292 AA; 32521 MW; F780D391984D66E3 CRC64;

Query Match 13.9%; Score 35; DB 11; Length 292;  
Best Local Similarity 100.0%; Pred. No. 1.6e-28;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 QHEIDLPTISLSDGEEPPPYQGPPCTQLQRDPEQQ 140  
Db 149 QHEIDLPTISLSDGEEPPPYQGPPCTQLQRDPEQQ 183

RESULT 6  
Q8BWJ4 PRELIMINARY; PRT; 306 AA.  
AC Q8BWJ4;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Protein C18ORF1 homolog.  
GN D18ERTD653E OR D330030L18RIK.

QY 3 ELEFVQIIIIIVVMVMVVTCLLSHYKLSARSFISRHSGRRREDALSS 53  
Db 34 ELEFVQIIIIIVVMVMVVTCLLSHYKLSARSFISRHSGRRREDALSS 84

RESULT 4  
Q9R5E2 PRELIMINARY; PRT; 200 AA.  
AC Q9R5E2;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Similar to chromosome 18 open reading frame 1 (Fragment).  
GN D18ERTD653E OR D330030L18RIK.

```
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=Heart;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK052338; BAC34943.1; -.
DR MGD; MGI:1277150; D18Erd653e.
DR InterPro; IPR002172; LDL_receptor_A.
DR Pfam; PF00057; ldl_recept_a; 1.
DR SMART; SM00192; LDLa; 1.
DR PROSITE; PS50068; LDLRA_2; 1.
SQ SEQUENCE 306 AA; 33857 MW; 043C2260AEE8729F CRC64;

Query Match 13.9%; Score 35; DB 11; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.7e-28;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 QHEIDLPTTISLSDGEEPPPYQGSPCTQLQRLDPEQQ 140
Db |||||
163 QHEIDLPTTISLSDGEEPPPYQGSPCTQLQRLDPEQQ 197

RESULT 7
Q89EU9 PRELIMINARY; PRT; 490 AA.
AC Q89EU9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE AprE protein.
GN APRE OR BLR6971.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamiasawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriiguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpou S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005960; BAC52236.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0015428; F:type I protein secretor activity; IEA.
DR GO; GO:0009306; P:protein secretion; IEA.
DR InterPro; IPR006143; HlyD.
DR InterPro; IPR003997; RtxD.
DR Pfam; PF00529; HlyD; 1.
DR PRINTS; PR01490; RTXTOXIND.
DR Complete proteome.
SQ SEQUENCE 490 AA; 54390 MW; AE3DF001D8F0F218 CRC64;

Query Match 3.2%; Score 8; DB 16; Length 490;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 232 APLESAAI 239
Db |||||
92 APLESAAI 99
```

```
RESULT 8
Q8CK70 PRELIMINARY; PRT; 1117 AA.
AC Q8CK70;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN SCO0072 OR SCJ4.53C OR SCJ11.01C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.P., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares S., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939104; CAD55416.1; -.
DR PIR; T36972; T36972.
DR InterPro; IPR006530; YD.
DR Pfam; PF05593; RHS repeat; 9.
DR TIGRFAMs; TIGR01643; YD_repeat_2x; 9.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1117 AA; 119294 MW; C3AD8A5834EFB867 CRC64;

Query Match 3.2%; Score 8; DB 16; Length 1117;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 ELNRESVR 149
Db |||||
646 ELNRESVR 653

RESULT 9
P94460 PRELIMINARY; PRT; 1271 AA.
AC P94460;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Peptide synthetase ORF5.
GN pps5.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=95227362; PubMed=7711903;
RA Tognoni A., Franchi E., Magistrelli C., Colombo E., Cosmina P.,
RA Grandi G.;
RT "A putative new peptide synthase operon in Bacillus subtilis: partial
RT characterization.";
RL Microbiology 141:645-648(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Grandi G.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
```



RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA de Ferra F., Tognoni A.;  
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z34883; CAAB4364.1; -.  
 DR HSSP; P14687; 1AMU.  
 DR GO; GO:0003824; P:catalytic activity; IEA.  
 DR GO; GO:0016788; P:hydrolase activity, acting on ester bonds; IEA.  
 DR GO; GO:0009058; P:biosynthesis; IEA.  
 DR InterPro; IPR000873; AMP-bind.  
 DR InterPro; IPR001242; Condensatn.  
 DR InterPro; IPR006162; Ppantne S.  
 DR InterPro; IPR006163; Pp\_bind.  
 DR InterPro; IPR001031; Thioesterase.  
 DR Pfam; PF00501; AMP-binding; 1.  
 DR Pfam; PF00668; Condensation; 1.  
 DR Pfam; PF00550; pp-binding; 1.  
 DR Pfam; PF00975; Thioesterase; 1.  
 DR PRINTS; PR00154; AMPBINDING.  
 DR PROSITE; PS00075; ACP DOMAIN; 1.  
 DR PROSITE; PS00455; AMP\_BINDING; 1.  
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.  
 KW Phosphopantetheine.  
 SQ SEQUENCE 1271 AA; 144057 MW; 1F8B115E0B936C72 CRC64;

Query Match 3.2%; Score 8; DB 2; Length 1271;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 KEKDKQKG 249  
 DB 111 KEKDKQKG 118  
 |||||

## RESULT 10

O31827  
 ID O31827 PRELIMINARY; PRT; 1279 AA.

AC O31827;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Peptide synthetase.

GN PPSE.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
 RA Guiseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,

RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,  
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*."  
 RL Nature 390:249-256(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 8-1279 FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98048468; PubMed=9387222;  
 RA Tosato V., Albertini A.M., Zotti M., Sonda S., Bruschi C.V.;  
 RT "Sequence completion, identification and definition of the fengycin  
 operon in *Bacillus subtilis* 168.";  
 RL Microbiology 143:3443-3450(1997).  
 DR EMBL; Z99113; CAB13713.1; -.  
 DR EMBL; Y13917; CAA74213.1; -.  
 DR PIR; E69681; E69681.  
 DR HSSP; P14687; 1AMU.  
 DR GO; GO:0003824; P:catalytic activity; IEA.  
 DR GO; GO:0016788; P:hydrolase activity, acting on ester bonds; IEA.  
 DR GO; GO:0009058; P:biosynthesis; IEA.  
 DR InterPro; IPR000873; AMP-bind.  
 DR InterPro; IPR001242; Condensatn.  
 DR InterPro; IPR006162; Ppantne S.  
 DR InterPro; IPR006163; Pp\_bind.  
 DR InterPro; IPR001031; Thioesterase.  
 DR Pfam; PF00501; AMP-binding; 1.  
 DR Pfam; PF00668; Condensation; 1.  
 DR Pfam; PF00550; pp-binding; 1.  
 DR Pfam; PF00975; Thioesterase; 1.  
 DR PRINTS; PR00154; AMPBINDING.  
 DR PROSITE; PS00075; ACP DOMAIN; 1.  
 DR PROSITE; PS00455; AMP\_BINDING; 1.  
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.  
 KW Phosphopantetheine; Complete proteome.  
 SQ SEQUENCE 1279 AA; 144618 MW; 22EA638DEFOCEBE0 CRC64;

Query Match 3.2%; Score 8; DB 16; Length 1279;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 KEKDKQKG 249  
 DB 118 KEKDKQKG 125  
 |||||

## RESULT 11

O8F917  
 ID O8F917 PRELIMINARY; PRT; 51 AA.

AC O8F917;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Hypothetical protein.

GN LA0208.  
 OS Leptospira interrogans.  
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.  
 OX NCBI\_TaxID=173;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;  
 RA Ren S.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE011210; AAN47407.1; -.  
 KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 51 AA; 6014 MW; 27D5715FF73D1BAF CRC64;  
Query Match 2.8%; Score 7; DB 16; Length 51;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ALEFPVQ 8  
Db 27 ALEFPVQ 33

RESULT 12  
O97510 PRELIMINARY; PRT; 61 AA.  
AC O97510;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Syntaxin 3 homolog (Fragment).  
OS Sus scrofa domestica (domestic pig).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9825;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Thyroid;  
RA Witke A., Koch M., Wahl R., Haring H.-U.;  
RT "cDNA homologous to human Syntaxin 3."  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF004135; AAC79692.1; -.  
DR InterPro; IPR000727; T SNARE.  
DR Pfam; PF05739; SNARE; 1.  
FT NON TER 1  
SQ SEQUENCE 61 AA; 6865 MW; 4637D61BF80DF5AB CRC64;

Query Match 2.8%; Score 7; DB 6; Length 61;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 IIIIVV 15  
Db 37 IIIIVV 43

RESULT 13  
Q8VB64 PRELIMINARY; PRT; 73 AA.  
AC Q8VB64;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Wsv124 (WSSV180).  
OS White spot syndrome virus (WSSV).  
OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae.  
OX NCBI\_TaxID=92652;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21548311; PubMed=11689662;  
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;  
RT "Complete genome sequence of the shrimp white spot bacilliform virus."  
RL J. Virol. 75:11811-11820(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Taiwan;  
RX MEDLINE=20517548; PubMed=11062040;  
RA Tsai M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.,  
RA Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;  
RT "Identification and characterization of a shrimp white spot syndrome

RT virus (WSSV) gene that encodes a novel chimeric polypeptide of  
RT cellular-type thymidine kinase and thymidylate kinase."  
RL Virology 277:100-110(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Taiwan;  
RX MEDLINE=21844071; PubMed=11853398;  
RA Chen L.L., Leu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,  
RA Lo C.F., Kou G.H.;  
RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white  
RT spot syndrome virus and characterization of the motif important for  
RT targeting VP35 to the nuclei of transfected insect cells."  
RL Virology 293:44-53(2002).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Taiwan;  
RA Lo C.F., Kou G.H.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF332093; AAL33128.1; -.  
DR EMBL; AF440570; AAL89048.1; -.  
SQ SEQUENCE 73 AA; 8174 MW; A7D244DD1AC79C70 CRC64;

Query Match 2.8%; Score 7; DB 12; Length 73;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 171 CPPSSNS 177  
Db 5 CPPSSNS 11

RESULT 14  
Q86SR9 PRELIMINARY; PRT; 95 AA.  
AC Q86SR9;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Similar to TRAF-binding protein domain.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon, Kidney, and Stomach;  
RA Strausberg R.;  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC048281; AAH48281.1; -.  
SQ SEQUENCE 95 AA; 11302 MW; 4C01331F90E1A45C CRC64;

Query Match 2.8%; Score 7; DB 4; Length 95;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 116 SLSDGEE 122  
Db 82 SLSDGEE 88

RESULT 15  
Q8LRN2 PRELIMINARY; PRT; 96 AA.  
AC Q8LRN2;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Putative class II small heat shock protein (Fragment).  
OS Ginkgo biloba (Ginkgo).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Ginkgophyta; Ginkgoales; Ginkgoaceae; Ginkgo.  
OX NCBI\_TaxID=3311;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Ayadi R., Papon N., Chenieux J.-C., Rideau M.,  
RA Tremouillaux-Guiller J.;  
RT "Cloning, characterization and expression of a partial cDNA encoding a  
RT low molecular weight heat shock protein in Ginkgo biloba L.";  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF507963; AAM34241.1; -.  
DR GO; GO:0003773; F:heat shock protein activity; IEA.  
DR InterPro; IPR002068; Hsp20.  
DR InterPro; IPR008978; HSP20\_chap.  
DR Pfam; PF00011; HSP20; 1.  
DR PROSITE; PS01031; HSP20; 1.  
KW Heat shock.  
FT NON TER 1 1  
SQ SEQUENCE 96 AA; 10755 MW; 25B61BB503A03547 CRC64;  
  
Query Match 2.8%; Score 7; DB 10; Length 96;  
Best Local Similarity 100.0%; Pred.No. 46;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 178 GISATCY 184  
|||  
Db 63 GISATCY 69

Search completed: May 20, 2004, 15:36:16  
Job time : 48 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 14:55:02 ; Search time 6855 Seconds  
(without alignments)  
6708.528 Million cell updates/sec

Title: US-09-857-826B-44  
Perfect score: 1061  
Sequence: 1 tcctccttggttcgggtga.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rod:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1057	99.6	1085	6	AX775887	AX775887 Sequence
2	961	90.6	969	6	BD272494	BD272494 Secreted
3	955	90.0	1913	6	BD272544	BD272544 Secreted
4	953	89.8	1061	9	BC015918	BC015918 Homo sapi
5	953	89.8	1383	6	AX775889	AX775889 Sequence
6	952	89.7	969	6	BD272514	BD272514 Secreted
7	926	87.3	1141	9	AF224278	AF224278 Homo sapi
8	912	86.0	1140	6	AR336830	AR336830 Sequence
9	910	85.8	969	6	BD272515	BD272515 Secreted
10	910	85.8	969	6	BD272516	BD272516 Secreted
11	903	85.1	1818	9	AY128643	AY128643 Homo sapi
12	902	85.0	4839	9	AF305616	AF305616 Homo sapi
13	756	71.3	756	6	BD272495	BD272495 Secreted
14	753	71.0	753	6	BD272534	BD272534 Secreted
15	749	70.6	1321	6	AX392417	AX392417 Sequence
16	708	66.7	759	6	BD272545	BD272545 Secreted
17	708	66.7	759	6	AR336831	AR336831 Sequence
18	701	66.1	861	6	AX392419	AX392419 Sequence
19	695	65.5	61505	9	AF305426	AF305426 Homo sapi
20	695	65.5	130435	9	HS718J7	AL035541 Human DNA
21	509	48.0	1583	6	AX593655	AX593655 Sequence
22	350	33.0	408	6	AX071267	AX071267 Sequence
23	322	30.3	812	6	AX011709	AX011709 Sequence
24	322	30.3	812	6	BD226320	BD226320 Pancreat
25	313	29.5	693	6	AX392430	AX392430 Sequence
26	157	14.8	150224	9	HSJ1059L7	AL121913 Human DNA
27	79	7.4	2570	9	AK056098	AK056098 Homo sapi
28	51	4.8	51	6	AX199565	AX199565 Sequence
29	44	4.1	411	6	BD272547	BD272547 Secreted
30	44	4.1	484	6	BD272546	BD272546 Secreted
31	44	4.1	648	6	BD272535	BD272535 Secreted
32	44	4.1	651	6	BD272505	BD272505 Secreted
33	44	4.1	651	10	AF220208	AF220208 Mus muscu
34	44	4.1	878	6	AX392428	AX392428 Sequence
35	44	4.1	1379	10	BC036995	BC036995 Mus muscu
36	44	4.1	1713	6	BD272504	BD272504 Secreted
37	44	4.1	1713	6	BD272517	BD272517 Secreted
38	44	4.1	1713	6	BD272518	BD272518 Secreted
39	44	4.1	1713	6	BD272519	BD272519 Secreted
40	44	4.1	156698	10	AL837509	AL837509 Mouse DNA
41	44	4.1	175754	2	AC110189	AC110189 Mus muscu
42	44	4.1	231930	2	AC134911	AC134911 Mus muscu
43	42	4.0	673	6	AX525744	AX525744 Sequence
44	42	4.0	249554	2	AC139417	AC139417 Rattus no
45	42	4.0	258632	2	AC111878	AC111878 Rattus no

ALIGNMENTS

RESULT 1  
AX775887  
LOCUS AX775887  
DEFINITION Sequence 157 from Patent WO03048202.  
ACCESSION AX775887  
VERSION AX775887.1 GI:32693605  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS Matsuda, A. and Muramatsu, S.  
TITLE NF-kB activating gene  
JOURNAL Patent: WO 03048202-A 157 12-JUN-2003;

AX775887 1085 bp mRNA linear PAT 14-JUL-2003





QY	281	GAGAGCACAGTGTTCAGGCAACCGAATCCAGAGCCGCGCAGGTCTACGCCCGCCCTCGGCC	340
DB	186	GAGAGCACAGTGTTCAGGCAACCGAATCCAGAGCCGCGCAGGTCTACGCCCGCCCTCGGCC	245
QY	341	ACCGACCGCTGGCCGTGCGCCCTTCGCCCCAGCGGGAGCGCTTCCACCGCTTCCAGGCC	400
DB	246	ACCGACCGCTGGCCGTGCGCCCTTCGCCCCAGCGGGAGCGCTTCCACCGCTTCCAGGCC	305
QY	401	ACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCCAACCATCTCGCTGTCCAGCGG	460
DB	306	ACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCCAACCATCTCGCTGTCCAGCGG	365
QY	461	GAGGAGCCCCACCCCTACCGAGCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAG	520
DB	366	GAGGAGCCCCACCCCTACCGAGCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAG	425
QY	521	CTGGAACTGAACCGGAGTCGCTGCGCGCACCCCAACAGAACCATCTTCGACAGTGAC	580
DB	426	CTGGAACTGAACCGGAGTCGCTGCGCGCACCCCAACAGAACCATCTTCGACAGTGAC	485
QY	581	CTGATGGATAGTCCAGGCTGGGCGGCCCTGCCCCCCAGCAGTAACCTCGGGCATCAGC	640
DB	486	CTGATGGATAGTCCAGGCTGGGCGGCCCTGCCCCCCAGCAGTAACCTCGGGCATCAGC	545
QY	641	GCCACGTGCTACGGCAGCGGGCGCGCATGGAGGGCGCGGCCACCTACAGCGAGTTC	700
DB	546	GCCACGTGCTACGGCAGCGGGCGCGCATGGAGGGCGCGGCCACCTACAGCGAGTTC	605
QY	701	ATCGGCCACTACCGGGTCTCCTTCCAGCACCATGGAGGGCGCGGCCACCTACAGCGAGTTC	760
DB	606	ATCGGCCACTACCGGGTCTCCTTCCAGCACCATGGAGGGCGCGGCCACCTACAGCGAGTTC	665
QY	761	CTGGAGGGGACCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGG	820
DB	666	CTGGAGGGGACCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGG	725
QY	821	AGCAAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCAGGGGGCGCGGCT	880
DB	726	AGCAAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCAGGGGGCGCGGCT	785
QY	881	GGGGCTCGTAGGTGAAGAGGAGAACCTCCCGCGCTTCTTAGAAGAGGAGTGAGAGGAA	940
DB	786	GGGGCTCGTAGGTGAAGAGGAGAACCTCCCGCGCTTCTTAGAAGAGGAGTGAGAGGAA	845
QY	941	GGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAATA	1000
DB	846	GGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAATA	905
QY	1001	TTTACATGTGATGTCTGCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA	1060
DB	906	TTTACATGTGATGTCTGCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA	965
QY	1061	A 1061	
DB	966	A 966	
RESULT 3			
LOCUS	BD272544	1913 bp	DNA linear
DEFINITION	Secreted proteins and nucleic acids encoding them.		
ACCESSION	BD272544		
VERSION	BD272544.1	GI:33082312	
KEYWORDS	JP 2002539773-A/53.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 1913)		
JOURNAL	Barnes, T.M., Holtzman, D.A., Sharp, J.D. and Fraser, C.C.		
	Secreted proteins and nucleic acids encoding them		
	Patent: JP 2002539773-A 53 26-NOV-2002.		

COMMENT	MILLENNIUM PHARMACEUTICALS INC
OS	Homo sapiens (human)
PN	JP 2002539773-A/53
PD	26-NOV-2002
PF	01-MAR-2000 JP 2000602247
PR	01-MAR-1999 US 60/122458
PI	THOMAS M BARNES, DOUGLAS A HOLTZMAN, JOHN D SHARP, CHRISTOPHER C
PI	FRASER
PC	C12N15/09, C07K14/47, C07K16/18, C12N5/10, C12P21/02, C12Q1/68, PC
	G01N33/15,
PC	G01N33/50, G01N33/53, G01N33/53, G01N33/53, G01N33/566, C12N15/00, C12N5/00 CC
	Secreted proteins and nucleic acids encoding them FH Key
	Location/Qualifiers
FT	source 1..1913
FT	/organism='Homo sapiens (human)'
FEATURES	Location/Qualifiers
source	1..1913
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
ORIGIN	
	Query Match 90.0%; Score 955; DB 6; Length 1913;
	Best Local Similarity 99.8%; Pred. No. 0;
	Matches 1055; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY	1 TCCTCCTTGGGTTGGGTGAAAGCGCTTCCTGCGAAACGAGCAATGGCGGAGCTGGAGTTGT 60
Db	
46	TCCTCCTTGGGTTGGGTGAAAGCGCTTCCTGCGAAACGAGCAATGGCGGAGCTGGAGTTGT 105
QY	61 TGGAGAACTGAAGCGGACGGTCTCTTCGAAACGAGCAATGGCGGAGCTGGAGTTGT 120
Db	
106	TGGAGAACTGAAGCGGACGGTCTCTTCGAAACGAGCAATGGCGGAGCTGGAGTTGT 165
QY	121 TCAGATCATCATCATCGTGGTGGTGATGATGGTGTGATGGTGTGATCATCATGGCTGCT 180
Db	
166	TCAGATCATCATCATCGTGGTGGTGATGATGGTGTGATGGTGTGATCATCATGGCTGCT 225
QY	181 GAGCCACTACAAGCTGTCTGCAACGGTCTTCATCAGCCGGCACAGCCAGGGCGGAGGAG 240
Db	
226	GAGCCACTACAAGCTGTCTGCAACGGTCTTCATCAGCCGGCACAGCCAGGGCGGAGGAG 285
QY	241 AGAAGATGCCCTGTCTCTCAGAAGGATGCTGTGGCCCTCGGAGAGCACAGTGTTCAGGCAA 300
Db	
286	AGAAGATGCCCTGTCTCTCAGAAGGATGCTGTGGCCCTCGGAGAGCACAGTGTTCAGGCAA 345
QY	301 CGGAATCCAGAGCCGAGGTCTACGCCCGCCCTCGGCCCAACCGACCGCTGGCCGTGCC 360
Db	
346	CGGAATCCAGAGCCGAGGTCTACGCCCGCCCTCGGCCCAACCGACCGCTGGCCGTGCC 405
QY	361 GCCCTTCGCCCCAGCGGAGCGCTTCACCGCTTCAGCCCAACCTATCCGTACCTGCAGCA 420
Db	
406	GCCCTTCGCCCCAGCGGAGCGCTTCACCGCTTCAGCCCAACCTATCCGTACCTGCAGCA 465
QY	421 CGAGATCGACCTGCCGCCCAACCATCTCGCTGTGACGCGGAGGAGCCCAACCTATCCCA 480
Db	
466	CGAGATCGACCTGCCGCCCAACCATCTCGCTGTGACGCGGAGGAGCCCAACCTATCCCA 525
QY	481 GGGCCCTTCACCCCTCCAGCTTCGGGACCCCGAGCAGCTGGAACCTGAACCGGGAGTC 540
Db	
526	GGGCCCTTCACCCCTCCAGCTTCGGGACCCCGAGCAGCTGGAACCTGAACCGGGAGTC 585
QY	541 GGTGCGGCACCCCAACAGAACCATCTTCGACAGTGACCTGATGGATAGTCCAGGCT 600
Db	
586	GGTGCAGCACCCCAACAGAACCATCTTCGACAGTGACCTGATGGATAGTCCAGGCT 645
QY	601 GGGCGGCCCTGCCCCCCCCAGCAGTAACTCGGGCATCAGCGCCACGTGCTACGGCAGCGG 660
Db	
646	GGGCGGCCCTGCCCCCCCCAGCAGTAACTCGGGCATCAGCGCCACGTGCTACGGCAGCGG 705
QY	661 CGGCGCATGGAGGGCGCGGCCCACTACAGCGAGGTTCATCGGCCACTACCCGGGGTC 720
Db	
706	CGGCGCATGGAGGGCGCGGCCCACTACAGCGAGGTTCATCGGCCACTACCCGGGGTC 765



QY 721 CTCCTTCCAGCACCAGCAGCAGTGGGGCCCTCTTGTGGAGGGGACCCGGCTCCA 780  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
766 CTCCTTCCAGCACCAGCAGCAGTGGGGCCCTCTTGTGGAGGGGACCCGGCTCCA 825  
QY 781 CCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGGATAAACA 840  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
826 CCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGGATAAACA 885  
QY 841 GAAAGGACACCCCTCTCTAGGTTCCAGGGGGCCGGCTGGGGCTGCGTAGGTGAAAG 900  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
886 GAAAGGACACCCCTCTCTAGGTTCCAGGGGGCCGGCTGGGGCTGCGTAGGTGAAAG 945  
QY 901 GCAGAACACTCCGGCTTCTTAGAAGAGGAGTGAGAGGAGCGGGGGCGCAGCAACGC 960  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
946 GCAGAACACTCCGGCTTCTTAGAAGAGGAGTGAGAGGAGCGGGGGCGCAGCAACGC 1005  
QY 961 ATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATTTACATGTGTCGTGTC 1020  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1006 ATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATTTACATGTGTCGTGTC 1065  
QY 1021 TGAATGCACAGCTAAGAGAGCTTGCAAAAAA 1057  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1066 TGAATGCACAGCTAAGAGAGCTTGCAAAAAA 1102

RESULT 4  
BC015918 1061 bp mRNA linear PRI 07-OCT-2003  
LOCUS Homo sapiens transmembrane, prostate androgen induced RNA, mRNA  
DEFINITION (CDNA clone MGC:20374 IMAGE:4559576), complete cds.  
ACCESSION BC015918  
VERSION BC015918.1 GI:16198474  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1061)

REFERENCE 1 (bases 1 to 1061)  
AUTHORS Klausner,R.D., Collins,P.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,  
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
Hopkins,R.P., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P.,  
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
Scheetz,T.E., Brownstein,M.J., Urdin,T.B., Toshiyuki,S.,  
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,  
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,  
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,  
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,  
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallos,D.E.,  
Schnerch,A., Schein,J.B., Jones,S.J. and Marra,M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences

TITLE  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
MEDLINE 22388257  
PUBMED 12477932  
REFERENCE 2 (bases 1 to 1061)  
AUTHORS Strausberg,R.  
TITLE Direct Submission  
JOURNAL Submitted (15-OCT-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline  
Schein, Duane Smallos, Michael Smith, Lorraine Spence, Jeff Stott,  
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 29 Row: e Column: 5  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Similarity but not  
identity to protein.

FEATURES  
source Location/Qualifiers  
1..1061  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:20374 IMAGE:4559576"  
/tissue\_type="Kidney, renal cell adenocarcinoma"  
/clone\_lib="NIH MGC 14"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
1..1061  
/gene="TMEPAI"  
/note="synonym: PMEPA1"  
/db\_xref="LocusID:56937"  
/db\_xref="MIM:606564"  
131..844  
/codon\_start=1  
/product="TMEPAI protein"  
/protein\_id="AAH15918.1"  
/db\_xref="GI:16198475"  
/db\_xref="LocusID:56937"  
/translation="MMVMVVITCLSHYKLSARSPISRHSQRRRREDALSSGCLMP  
SESTVSGNGIPQVYAPPRDLAVPPAQRERFHRFQPTYPYLOHIDLPPTISL  
SDGEBPPYQGFCTQLRDPQQLLEINRESVRAPNRTIFDSLDMSARLGGPCPPSS  
NSGISATCYGSGRMEGPPPTYSVIGHYPCSSFHQSSGPPSLLEGTRLHHTHIA  
LESAAIWSKERDQKHPL"

ORIGIN  
Query Match 89.8%; Score 953; DB 9; Length 1061;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 953; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 105 CGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGTGGTGGTGGTGGTGG 164  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
90 CGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGTGGTGGTGGTGGTGG 149  
QY 165 TGATCAGTGCCTGCTGAGCCACTACAAGCTGTCTGCAAGGCTCTTCATCAGCCGCGACA 224  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
150 TGATCAGTGCCTGCTGAGCCACTACAAGCTGTCTGCAAGGCTCTTCATCAGCCGCGACA 209  
QY 225 GCCAGGGGGGAGGAGAGATGCCCTGTCTCCTCAGAGGATGCCTGTGGCCCTCGGAGA 284  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
210 GCCAGGGGGGAGGAGAGATGCCCTGTCTCCTCAGAGGATGCCTGTGGCCCTCGGAGA 269  
QY 285 GCACAGTGTGAGCAACGGAATCCAGAGCCGCGAGTCTACGCCCGCTCGGCCACCG 344  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
270 GCACAGTGTGAGCAACGGAATCCAGAGCCGCGAGTCTACGCCCGCTCGGCCACCG 329  
QY 345 ACCGCTGCGCGTCCCTTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACCT 404  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
330 ACCGCTGCGCGTCCCTTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACCT 389  
QY 405 ATCGGTACCTGCAGCAGAGATCGACCTGCCGCCACCACCATCTCGCTGTTCAGACGGGGAGG 464

Db 390 ATCGTACCTGCAGCAGAGATCGACCTGCGCCACCATCTCGCTGTGACAGCGGGAGG 449  
Qy 465 AGCCCCCACCCTACCAGGGCCCTGACACCTTCCAGCTTCGGGACCCCGAGCAGCAGCTGG 524  
Db 450 AGCCCCCACCCTACCAGGGCCCTGACACCTTCCAGCTTCGGGACCCCGAGCAGCAGCTGG 509  
Qy 525 AACTGAACCGGAGTGGTGGCGGACACCCCAACAGAACCACTTTCGACAGTGCCTGA 584  
Db 510 AACTGAACCGGAGTGGTGGCGGACACCCCAACAGAACCACTTTCGACAGTGCCTGA 569  
Qy 585 TGGATAGTCCAGGCTGGGGCCCTGCCCCCCCCAGCAGTAACCTCGGGCATCAGCGCCA 644  
Db 570 TGGATAGTCCAGGCTGGGGCCCTGCCCCCCCCAGCAGTAACCTCGGGCATCAGCGCCA 629  
Qy 645 CTGTCTACGGCAGCGCGGCGCATGGAGGGCGCGCCGCCACCTACAGCGAGGTATCG 704  
Db 630 CTGTCTACGGCAGCGCGGCGCATGGAGGGCGCGCCGCCACCTACAGCGAGGTATCG 689  
Qy 705 GCCACTACCGGGGTCTCTTCAGACACCATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCA 824  
Db 750 AGGGGACCGGCTCCACACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCA 809  
Qy 825 AAGAGAAAGGATAAACAGAAAGGACACCCCTCTAGSGTCCCGAGGGGGCGGGCTGGG 884  
Db 810 AAGAGAAAGGATAAACAGAAAGGACACCCCTCTAGSGTCCCGAGGGGGCGGGCTGGG 869  
Qy 885 CTGCGTAGGTGAAGGCGAGAACACTCCGCGCTTCTTAGAGAGGAGTGAGAGGAGCGG 944  
Db 870 CTGCGTAGGTGAAGGCGAGAACACTCCGCGCTTCTTAGAGAGGAGTGAGAGGAGCGG 929  
Qy 945 GGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATTTA 1004  
Db 930 GGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATTTA 989  
Qy 1005 CATGTGATGTCTGGTCTGAATGCAAGCTAAGAGAGCTTTCGCAAAAAA 1057  
Db 990 CATGTGATGTCTGGTCTGAATGCAAGCTAAGAGAGCTTTCGCAAAAAA 1042

RESULT 5  
AX775889  
LOCUS AX775889 1383 bp mRNA linear PAT 14-JUL-2003  
DEFINITION Sequence 159 from Patent WO03048202.  
ACCESSION AX775889  
VERSION AX775889.1 GI:32693607  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Matsuda, A. and Muramatsu, S.  
TITLE NF-kB activating gene  
JOURNAL Patent: WO 03048202-A 159 12-JUN-2003;  
Asahi Kasei Kabushiki Kaisha (JP)  
FEATURES  
source location/Qualifiers  
1..1383  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
321..1184  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAB11644.1"  
/db\_xref="GI:32693608"  
/translation="MRLMGVNSTAAAGQPNVSTCNKRLSFQSMETLEFVQI  
IIIVVMVMVVIITCLSHYKLSARSPISRHSQRRREDALSSGGLMPSESTVSGN  
GIPEQVYAPPRPTDLAVPPAQRERFHRFQPTPYLQHEIDLPTTSLSDGSEPPP

YQGPCTQLRDPEQQLBNRESVRAPPNRTIFDSDLMDLSARLGSPCPSPSSNSGISATC  
YSGGRMEGPPPTTYSEVIGHYPGSSPFHQSSGPPSLLEGTLHHTHIAPLESAIWS  
KEKDQKQKGPL"  
ORIGIN  
Query Match 89.8%; Score 953; DB 6; Length 1383;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 953; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 105 CGGAGCTGGAGTTTGTTCAGATCATCATCTGTTGGTGTGATGATGTTGGTGGTGG 164  
Db 430 CGGAGCTGGAGTTTGTTCAGATCATCATCTGTTGGTGTGATGATGTTGGTGGTGG 489  
Qy 165 TGATCAGCTGCCTGCTGAGGCCACTACAAAGCTGTCTGACCGTCTTCTATCAGCCGACCA 224  
Db 490 TGATCAGCTGCCTGCTGAGGCCACTACAAAGCTGTCTGACCGTCTTCTATCAGCCGACCA 549  
Qy 225 GCCAGGGCGGAGAGAGAGATGCTCTCTCAGAAAGATGCTCTGTGGCCCTCGGAGA 284  
Db 550 GCCAGGGCGGAGAGAGAGATGCTCTCTCAGAAAGATGCTCTGTGGCCCTCGGAGA 609  
Qy 285 GCACAGTGTGAGGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCTCGGCCCCACCG 344  
Db 610 GCACAGTGTGAGGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCTCGGCCCCACCG 669  
Qy 345 ACCGCTTGGCGCTGCGCCCTTTCGCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCCACT 404  
Db 670 ACCGCTTGGCGCTGCGCCCTTTCGCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCCACT 729  
Qy 405 ATCCGTACCTGCGAGCAAGAGATGACCTGCGCCCGCCACCATCTCGCTGTGACAGCGGGAGG 464  
Db 730 ATCCGTACCTGCGAGCAAGAGATGACCTGCGCCCGCCACCATCTCGCTGTGACAGCGGGAGG 789  
Qy 465 AGCCCCCACCCTACAGGGCCCTGCGACCTTCCAGCTTCGGGACCCCGAGCAGCAGCTGG 524  
Db 790 AGCCCCCACCCTACAGGGCCCTTGCACCTTCCAGCTTCCGGGACCCCGAGCAGCAGCTGG 849  
Qy 525 AACTGAACCGGGAGTGGTGGCGGACCCCGCCCAACAGAACCATCTTTCGACAGTGCCTGA 584  
Db 850 AACTGAACCGGGAGTGGTGGCGGACCCCGCCCAACAGAACCATCTTTCGACAGTGCCTGA 909  
Qy 585 TGGATAGTCCAGGCTGGGGCGCCCTGCGCCCGCCCGAGCAGTAACCTCGGGCATCAGCGCCA 644  
Db 910 TGGATAGTCCAGGCTGGGGCGCCCTGCGCCCGCCCGAGCAGTAACCTCGGGCATCAGCGCCA 969  
Qy 645 CGTGTACCGGCGGCGGCGCATGGAGGGCGCGCCCGCCCATACAGCGAGTGCCTGA 704  
Db 970 CGTGTACCGGCGGCGGCGCATGGAGGGCGCGCCCGCCCATACAGCGAGTGCCTGA 1029  
Qy 705 GCCACTACCGGGGTCTCTTCCAGCACCAGCAGCAGTGGGCGCCCTCTCTGCTGG 764  
Db 1030 GCCACTACCGGGGTCTCTTCCAGCACCAGCAGCAGTGGGCGCCCTCTCTGCTGG 1089  
Qy 765 AGGGGACCGGCTCCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCA 824  
Db 1090 AGGGGACCGGCTCCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCA 1149  
Qy 825 AAGAGAAAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGG 884  
Db 1150 AAGAGAAAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGG 1209  
Qy 885 CTGCGTAGGTGAAGGCGCAGAACACTCGCGCTTCTTAGAGAGGAGTGAGAGGAGCGG 944  
Db 1210 CTGCGTAGGTGAAGGCGCAGAACACTCGCGCTTCTTAGAGAGGAGTGAGAGGAGCGG 1269  
Qy 945 GGGGGCGCAGCAACGCATCGTGTGGCCCTTCCCTCCCGCTCCCTCTCTGTATAAATTTA 1004  
Db 1270 GGGGGCGCAGCAACGCATCGTGTGGCCCTTCCCTCCCGCTCCCTCTCTGTATAAATTTA 1329  
Qy 1005 CATGTGATGTCTGGTCTGAATGCAAGCTAAGAGAGCTTTCGCAAAAAA 1057  
Db 1330 CATGTGATGTCTGGTCTGAATGCAAGCTAAGAGAGCTTTCGCAAAAAA 1382





glandular epithelial cells; similar to the predicted protein encoded by sequence deposited at GenBank Accession Number NP\_004329"  
/codon\_start=1  
/product="PMEPA1 protein"  
/protein\_id="AAF86322.1"  
/db\_xref="GI:9255809"  
/translation="MARLEFVOIIIIIVVMVMVVVITCLLSHYKLSRSPISRHSQGRRREDALSSGCLMPSESTVSGNGIPPEQVYAPPRPTDLRAVPPPAQERFHRFPQTYPYLQHEIDLPTISLSDGEPPPYQGPCTQLQRDPQQLNLRNRESVRAPNRITFDSDLMSARLGGPCPSSNSGISATCYGSGGRMBGPPTYSEVIGHYPGSSFQHQQSSGPPSLLEGTRLHHTHIAPLESAAIWSKERDKQKHPL"

ORIGIN

Query Match		87.3%;	Score 926;	DB 9;	Length 1141;
Best Local Similarity		99.9%;	Pred. No. 0;		
Matches 976;		Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	81	GTCTCCTCGAAACCCAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGG	140		
Db	76	GTCTCCTCGAAACCCAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGG	135		
QY	141	TGGTGATGATGGTGGTGGTGATCACTGCTGCTGAGCCACTACAAGCTGTCTG	200		
Db	136	TGGTGATGATGGTGGTGGTGATCACTGCTGCTGAGCCACTACAAGCTGTCTG	195		
QY	201	CACGGTCTTTCATCAGCCGACAGCCAGCGGGCGGAGAGAGAAGATGCCCTGTCTCAG	260		
Db	196	CACGGTCTTTCATCAGCCGACAGCCAGCGGGCGGAGAGAGAAGATGCCCTGTCTCAG	255		
QY	261	AAGGATGCCCTGTGGCCCTCGGAGAGCACAGTGTAGGCAACGGAATCCAGAGCCGAGG	320		
Db	256	AAGGATGCCCTGTGGCCCTCGGAGAGCACAGTGTAGGCAACGGAATCCAGAGCCGAGG	315		
QY	321	TCTACGCCCGCCTCGGCCACCGACCGCCTGGCCGTGCGCCCTTGGCCAGCGGAGC	380		
Db	316	TCTACGCCCGCCTCGGCCACCGACCGCCTGGCCGTGCGCCCTTGGCCAGCGGAGC	375		
QY	381	GCTTCCACCGCTTCCAGCCCACTATCCGTACCTGACGACAGAGATGACCTGCGGCCCA	440		
Db	376	GCTTCCACCGCTTCCAGCCCACTATCCGTACCTGACGACAGAGATGACCTGCGGCCCA	435		
QY	441	CCATCTCGTGTAGACGGGAGGAGCCCCACCTACCGAGGGCCCCCTGCACCTCCAGC	500		
Db	436	CCATCTCGTGTAGACGGGAGGAGCCCCACCTACCGAGGGCCCCCTGCACCTCCAGC	495		
QY	501	TTCCGGACCCCGAGCAGCTGGAACCGGAGTGGTGGCGGCCCCCTGCACCCCAACA	560		
Db	496	TTCCGGACCCCGAGCAGCTGGAACCGGAGTGGTGGCGGCCCCCTGCACCCCAACA	555		
QY	561	GAACCATCTTCAGACAGTACCTGATGATAGTGCAGAGCTGGGCGGCCCTTGCACCCCA	620		
Db	556	GAACCATCTTCAGACAGTACCTGATGATAGTGCAGAGCTGGGCGGCCCTTGCACCCCA	615		
QY	621	GCAGTAACTCGGGCATCAGCGCCACGTCACGCGAGCGCGGGCGCATGGAGGGCGGC	680		
Db	616	GCAGTAACTCGGGCATCAGCGCCACGTCACGCGAGCGCGGGCGCATGGAGGGCGGC	675		
QY	681	CGCCCACTACAGCGAGGTTCATCGGCCACTACCCGGGTCTCTCTCCAGCACAGCAGA	740		
Db	676	CGCCCACTACAGCGAGGTTCATCGGCCACTACCCGGGTCTCTCTCCAGCACAGCAGA	735		
QY	741	GCAGTGGGCGCCCTCTTGTGGAGGGGACCCGGCTCCACACACACATCGCGCCCC	800		
Db	736	GCAGTGGGCGCCCTCTTGTGGAGGGGACCCGGCTCCACACACACATCGCGCCCC	795		
QY	801	TAGAGAGCGCAGCCATCTGGAGCAAGAGAGAGATAAACAGAAAGGACACCTCTTAGG	860		
Db	796	TAGAGAGCGCAGCCATCTGGAGCAAGAGAGAGATAAACAGAAAGGACACCTCTTAGG	855		
QY	861	GTCCCCAGGGGGCGCGCTGGGGCTGCGTAGGTGAAAGGAGCAACACTCGCGCTTCT	920		
Db	856	GTCCCCAGGGGGCGCGCTGGGGCTGCGTAGGTGAAAGGAGCAACACTCGCGCTTCT	915		

RESULT 8					
AR336830					
LOCUS		AR336830	1140 bp	DNA	linear
DEFINITION		Sequence 1 from patent US 6566130.			
ACCESSION		AR336830			
VERSION		AR336830.1 GI:33722680			
KEYWORDS		Unknown.			
SOURCE		Unknown.			
ORGANISM		Unclassified.			
REFERENCE		1 (bases 1 to 1140)			
AUTHORS		Srivastava, S., Moul, J. W., Xu, L. L. and Segawa, T.			
TITLE		Androgen-regulated gene expressed in prostate tissue			
JOURNAL		Patent: US 6566130-A 1 20-MAY-2003;			
FEATURES		Location/Qualifiers			
		source			
		1..1140			
		/organism="unknown"			
		/mol_type="genomic DNA"			
ORIGIN					

Query Match 86.0%; Score 912; DB 6; Length 1140;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 962; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	95	CAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGTGGTGGTGGTG	154		
Db	89	CAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGTGGTGGTGGTG	148		
QY	155	ATGGTGGTGGTGATCAGTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTTCATC	214		
Db	149	ATGGTGGTGGTGATCAGTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTTCATC	208		
QY	215	AGCCGGCACAGCCAGGGCGGAGGAGAGAGATGCCCTGTCTCAGAAAGGATGCCTGTGG	274		
Db	209	AGCCGGCACAGCCAGGGCGGAGGAGAGATGCCCTGTCTCAGAAAGGATGCCTGTGG	268		
QY	275	CCCTCGGAGAGCACAGTGTGAGCAACGGGAATCCAGAGCCGCGAGGTCTACGCCCGCCT	334		
Db	269	CCCTCGGAGAGCACAGTGTGAGCAACGGGAATCCAGAGCCGCGAGGTCTACGCCCGCCT	328		
QY	335	CGGCCACCGACCGCCTGGCCGTGCGGCCCTTGGCCAGGGGAGCGCTTCCACCGCTTC	394		
Db	329	CGGCCACCGACCGCCTGGCCGTGCGGCCCTTGGCCAGGGGAGCGCTTCCACCGCTTC	388		
QY	395	CAGCCCACTATCCGTACCTGACGACGAGATCGACCTGCGGCCCAACCATCTCGCTGTCA	454		
Db	389	CAGCCCACTATCCGTACCTGACGACGAGATCGACCTGCGGCCCAACCATCTCGCTGTCA	448		
QY	455	GAAGGGGAGGAGCCCCACCCCTACAGGGGCCCTGCACCCCTCCAGCTTCGGGACCCCGAG	514		
Db	449	GAAGGGGAGGAGCCCCACCCCTACAGGGGCCCTGCACCCCTCCAGCTTCGGGACCCCGAG	508		
QY	515	CAGCAGCTGGAACCTGAACCGGGAGTGGTGGCGGCAACCCCAACAGAACCATCTTCGAC	574		
Db	509	CAGCAGCTGGAACCTGAACCGGGAGTGGTGGCGGCAACCCCAACAGAACCATCTTCGAC	568		
QY	575	AGTGACCTGATGGATAGTCCAGGGCTGGGGCGGCCCTTGCGGCCCCCGCAGCAGTAACCTCGGGC	634		
Db	569	AGTGACCTGATGGATAGTCCAGGGCTGGGGCGGCCCTTGCGGCCCCCGCAGCAGTAACCTCGGGC	628		



QY	635	ATCAGCGCCACGTGCTACGGCAGCGGGGGCGCATGGAGGGGCGCGCCACCTACAGC	694
DB	629	ATCAGCGCCACGTGCTACGGCAGCGGGGGCGCATGGAGGGGCGCGCCACCTACAGC	688
QY	695	GAGGTTCATCGGCCACTACCCCGGGGTCTCTCCAGCACACAGCAGAGCAGTGGGCGCGCC	754
DB	689	GAGGTTCATCGGCCACTACCCCGGGGTCTCTCTCCAGCACACAGCAGAGCAGTGGGCGCGCC	748
QY	755	TCCTTGCTGGAGGGACCCGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCC	814
DB	749	TCCTTGCTGGAGGGACCCGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCC	808
QY	815	ATCTGGAGCAAGAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGC	874
DB	809	ATCTGGAGCAAGAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGC	868
QY	875	CGGGCTGGGGCTGCGTAGGTGAAAAGGCAGAACACTCCGCGCTTCTTAGAAGAGGAGTGA	934
DB	869	CGGGCTGGGGCTGCGTAGGTGAAAAGGCAGAACACTCCGCGCTTCTTAGAAGAGGAGTGA	928
QY	935	GAGGAAGGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCCACCTCCCTGTGTA	994
DB	929	GAGGAAGGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCCACCTCCCTGTGTA	988
QY	995	TAAATATTTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA	1054
DB	989	TAAATATTTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA	1048
QY	1055	AAA 1057	
DB	1049	AAA 1051	

RESULT 9			
BD272515			
LOCUS	BD272515	969 bp	DNA linear
DEFINITION	Secreted proteins and nucleic acids encoding them.		

ORGANISM: Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 969)  
Barnes, T.M., Holtzman, D.A., Sharp, J.D. and Fraser, C.C.  
Secreted proteins and nucleic acids encoding them  
Patent: JP 2002539773-A 24 26-NOV-2002;  
JOURNAL

## ORIGIN

```

Query Match      85.8%; Score 910; DB 6; Length 969;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 960: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY	101	ATGGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGCTGATGAATGATGGTG	160
DB			
QY	6	ATGGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGCTGATGAATGATGGTG	65
DB			
QY	161	GTGGTGATCACTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTTCATCAGCCGG	220
DB			
QY	66	GTGGTGATCACTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTTCATCAGCCGG	125
DB			
QY	221	CACAGCCAGGGCGGAGGAGAAAGATGCCCTGTCTCAGAAGGATGCCCTGTGGCCCTCG	280
DB			
QY	126	CACAGCCAGGGCGGAGGAGAACGATGCCCTGTCTCAGAAGGATGCCCTGTGGCCCTCG	185
DB			
QY	281	GAGAGCACAGTGTTCAGGCAACGGAAATCCCAGAGCCGCGAGGTCTACGCCCGCCTCGGCC	340
DB			
QY	186	GAGAGCACAGTGTTCAGGCAACGGAAATCCCAGAGCCGCGAGGTCTACGCCCGCCTCGGCC	245
DB			
QY	341	ACCAGCCGCTGGCCGCTGCGCCCTTCGCCCCAGGGGAGCGCTTCCACCGCTTCCAGCCCC	400
DB			
QY	246	ACCAGCCGCTGGCCGCTGCGCCCTTCGCCCCAGGGGAGCGCTTCCACCGCTTCCAGCCCC	305
DB			
QY	401	ACCTATCCGTACCTGTCAGCAACGAGATCGACCTGCCGCCACCATCTCGCTGTTCAGACGGG	460
DB			
QY	306	ACCTATCCGTACCTGTCAGCAACGAGATCGACCTGCCGCCACCATCTCGCTGTTCAGACGGG	365
DB			
QY	461	GAGGAGCCCGACCTTACCAGGGCCCTGTCACCTCCAGCTTCGGGACCCCGAGCAGCAG	520
DB			
QY	366	GAGGAGCCCGACCTTACCAGGGCCCTGTCACCTCCAGCTTCGGGACCCCGAGCAGCAG	425
DB			
QY	521	CTGGAACCTGAACCGGGAGTGGTGCGCGCACCCCAAACAGAACCATCTTCGACAGTGAC	580
DB			
QY	426	CTGGAACCTGAACCGGGAGTGGTGCGCGCACCCCAAACAGAACCATCTTCGACAGTGAC	485
DB			
QY	581	CTGATGGATAGTCCAGGCTGGGCGGCCCTTGCCCCCGCAGCAGTAACCTCGGGCATCAGC	640
DB			
QY	486	CTGATGGATAGTCCAGGCTGGGCGGCCCTTGCCCCCGCAGCAGTAACCTCGGGCATCAGC	545
DB			
QY	641	GCCACGTGTACGGCAGCGGGCGGCATGGAGGGGCGCGCCACCTACAGCGAGGTC	700
DB			
QY	546	GCCACGTGTACGGCAGCGGGCGGCATGGAGGGGCGCGCCACCTACAGCGAGGTC	605
DB			
QY	701	ATCGGCCACTACCGGGGTCTCTCTCCAGCACCCAGCAGCAGTGGGCGGCCCTCCTTG	760
DB			
QY	606	ATCGGCCACTACCGGGGTCTCTCTCCAGCACCCAGCAGCAGTGGGCGGCCCTCCTTG	665
DB			
QY	761	CTGGAGGGGACCGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGG	820
DB			
QY	666	CTGGAGGGGACCGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGG	725
DB			
QY	821	AGCAAGAGAGGATAAACAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCT	880
DB			
QY	726	AGCAAGAGAGGATAAACAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCT	785
DB			
QY	881	GGGGCTGCTAGGTGAAAAGGCGAGAACACTCGCGCTCTTAGAAGAGGAGTGAGAGGAA	940
DB			
QY	786	GGGGCTGCTAGGTGAAAAGGCGAGAACACTCGCGCTCTTAGAAGAGGAGTGAGAGGAA	845
DB			
QY	941	GGGGGGGGCGCAGCAACGATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAATA	1000
DB			
QY	846	GGGGGGGGCGCAGCAACGATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAATA	905
DB			
QY	1001	TTTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCACAAAAA	1060
DB			
QY	906	TTTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCACAAAAA	965
DB			
QY	1061	A 1061	
DB			
QY	966	A 966	

RESULT 10  
BD272516  
LOCUS















Qy	345	ACGCGCTGGCCGTCGCGCCCTTCGCCCCAGCGGAGCGCTTCCACCGCTTCCAGCCCCACCT	404
Db	762	ACGCGCTGGCCGTCGCGCCCTTCGCCCCAGCGGAGCGCTTCCACCGCTTCCAGCCCCACCT	821
Qy	405	ATCCGTACCTGACGACGAGATCGACCTGCGGCCACCATCTCGCTGTGACACGGGGAGG	464
Db	822	ATCCGTACCTGACGACGAGATCGACCTGCGGCCACCATCTCGCTGTGACACGGGGAGG	881
Qy	465	AGCCCCACCTTACAGGGCCCTTGCACTCCAGCTTCGGGACCCCGAGGACGAGCTGG	524
Db	882	AGCCCCACCTTACAGGGCCCTTGCACTCCAGCTTCGGGACCCCGAGGACGAGCTGG	941
Qy	525	AACGTGAACCGGAGTCGGTGGCGGCACCCCAAACAGAACCATCTTCGACAGTGAACCTGA	584
Db	942	AACGTGAACCGGAGTCGGTGGCGGCACCCCAAACAGAACCATCTTCGACAGTGAACCTGA	1001
Qy	585	TGATAGTGCCAGGCTGGGGGGCCCTTGCCCCCAGCAGTAACCTCGGGCATCAGCGCCA	644
Db	1002	TGATAGTGCCAGGCTGGGGGGCCCTTGCCCCCAGCAGTAACCTCGGGCATCAGCGCCA	1061
Qy	645	CGTGCTACGGCAGCGGGGGCGCATGGAGGGCGCGCCGCCACCTACAGCGAGGTCTATCG	704
Db	1062	CGTGCTACGGCAGCGGGGGCGCATGGAGGGCGCGCCGCCACCTACAGCGAGGTCTATCG	1121
Qy	705	GCCACTACCGGGGTCTCTCTTCCAGCACAGCAGAGCAGTGGGGCGCCCTCTCTTGTGG	764
Db	1122	GCCACTACCGGGGTCTCTCTTCCAGCACAGCAGAGCAGTGGGGCGCCCTCTCTTGTGG	1181
Qy	765	AGGGGACCGGCTCCACACACACATCGGGCCCTTAGAGAGCGCAGCCATCTGGAGCA	824
Db	1182	AGGGGACCGGCTCCACACACACATCGGGCCCTTAGAGAGCGCAGCCATCTGGAGCA	1241
Qy	825	AAGAGAGGATAAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGG	884
Db	1242	AAGAGAGGATAAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGG	1301
Qy	885	CTGCGTAGGTGAAAAGGCAG	904
Db	1302	CTGCGTAGGTGAAAAGGCAG	1321

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: May 26, 2004, 14:52:22 ; Search time 667 Seconds  
(without alignments)  
6757.628 Million cell updates/sec

Title: US-09-857-826B-44  
Perfect score: 1061  
Sequence: 1 tcctccttggttggtga.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: OLIGO\_NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002s:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1061	100.0	1061	3 AAA47429	AAA47429 Sequence
2	1057	99.6	1085	9 ADC37324	ADC37324 Nuclear f
3	1056	99.5	1334	7 ABZ36103	ABZ36103 Human sec
4	961	90.6	969	3 AAA75151	AAA75151 CDNA enco
5	953	89.8	1383	9 ADC37326	ADC37326 Nuclear f
6	952	89.7	969	3 AAA75163	AAA75163 CDNA clon
7	926	87.3	1140	6 ABK92120	ABK92120 Prostate
8	926	87.3	1141	9 ADB75588	ADB75588 Prostate
9	926	87.3	1850	7 ACC49536	ACC49536 Tumour-as
10	912	86.0	1140	9 AAD60105	AAD60105 Human and
11	910	85.8	969	3 AAA75164	AAA75164 CDNA clon
12	910	85.8	969	3 AAA75165	AAA75165 CDNA clon
13	902	85.0	4839	7 ACC49552	ACC49552 Tumour-as
14	891	84.0	1066	4 AAI57868	AAI57868 Human pol
15	788	74.3	806	7 ACC49537	ACC49537 Tumour-as
16	749	70.6	1321	6 ABK12137	ABK12137 Human cDN
17	548	51.6	1069	4 AAI59654	AAI59654 Human pol
18	509	48.0	1583	6 ABS61424	ABS61424 Prostate
19	415	39.1	474	7 ABZ84732	ABZ84732 Toxicolog
20	350	33.0	408	5 AAF65983	AAF65983 Novel hum
21	322	30.3	812	2 AAZ52964	AAZ52964 Human pro
22	313	29.5	693	6 ABK12143	ABK12143 Human MIV
23	269	25.4	467	8 ACH14862	ACH14862 Human adu

24	229	21.6	254	3	AAA41265	AAA41265 Human sec
25	93	8.8	426	5	AAS84502	AAS84502 DNA encod
26	91	8.6	1879	5	AAS84503	AAS84503 DNA encod
27	68	6.4	522	6	ABT10027	ABT10027 Human bre
28	60	5.7	60	6	ABN40872	ABN40872 Human spl
29	51	4.8	51	4	AAH89714	AAH89714 Human cod
30	44	4.1	878	6	ABK12142	ABK12142 Mouse cDN
31	44	4.1	1713	3	AAA75167	AAA75167 cDNA clon
32	44	4.1	1713	3	AAA75166	AAA75166 cDNA clon
33	44	4.1	1713	3	AAA75152	AAA75152 cDNA enco
34	44	4.1	1713	3	AAA75168	AAA75168 cDNA clon
35	42	4.0	673	6	ABT09178	ABT09178 Phase-1 R
36	32	3.0	577	4	AAI23118	AAI23118 Probe #13
37	32	3.0	577	4	ABA68210	ABA68210 Human foe
38	32	3.0	577	4	AAI48425	AAI48425 Probe #17
39	32	3.0	577	4	ABA50270	ABA50270 Human bre
40	32	3.0	577	4	ABA35222	ABA35222 Probe #13
41	32	3.0	577	4	AAK42346	AAK42346 Human bon
42	32	3.0	577	4	AAK16592	AAK16592 Human bra
43	32	3.0	577	4	ABS41955	ABS41955 Human liv
44	32	3.0	577	5	AAI08772	AAI08772 Probe #87
45	32	3.0	577	6	ABS16402	ABS16402 Human gen

ALIGNMENTS

RESULT 1

AAA47429  
ID AAA47429 standard; DNA; 1061 BP.  
XX  
AC AAA47429;  
XX  
DT 20-OCT-2000 (first entry)  
XX  
DE Sequence encoding human neuron-associated protein.  
XX  
KW Neuron associated protein; NEUAP; neurological disorder; epilepsy;  
KW ischemic cerebrovascular disease; stroke; cerebral neoplasm;  
KW Alzheimer's disease; Pick's disease; Huntington's disease; dementia;  
KW Parkinson's disease; demyelinating disease; meningitis; prion disease;  
KW kuru; Creutzfeldt-Jakob disease; neurofibromatosis; cerebral palsy;  
KW muscular dystrophy; central nervous system; CNS;  
KW peripheral nervous system; PNS; myopathy; schizophrenia;  
KW actinic keratosis; arteriosclerosis; atherosclerosis; bursitis;  
KW cirrhosis; hepatitis; mixed connective tissue disease; MCTD;  
KW myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;  
KW autoimmune disease; inflammation; acquired immunodeficiency syndrome;  
KW AIDS; Addison's disease; adult respiratory distress syndrome; allergy;  
KW ankylosing spondylitis; amyloidosis; anaemia; asthma;  
KW Werner syndrome; trauma; human; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 101..859  
FT /\*tag= a  
FT /product= "Neuron associated protein"  
XX  
PN WO200034477-A2.  
XX  
PD 15-JUN-2000.  
XX  
PP 10-DEC-1999; 99WO-US030408.  
XX  
PR 11-DEC-1998; 98US-00210083.  
PR 09-FEB-1999; 99US-0119365P.  
PR 16-MAR-1999; 99US-0124687P.  
XX  
PA (INCY-) INCYTE PHARM INC.  
XX  
PI Tang YT, Yue H, Baughn MR, Hillman JL, Lal P, Au-Young J;  
PI Yang J, Lu DAM, Azimzai Y;



XX WPI; 2000-423423/36.  
DR P-PSDB; AAB01388.  
XX  
PT New human neuron-associated proteins and polynucleotides encoding them,  
PT useful for diagnosis, treatment and prevention of cell proliferative  
PT disorders including cancer, neuronal and neurological disorders.  
XX  
PS Claim 9; Page 136; 145pp; English.  
XX  
CC Human neuron-associated proteins (NEUAP) can be used for treating or  
CC preventing a disorder associated with decreased expression or activity of  
CC NEUAP. Antagonists of NEUAP are useful for treating or preventing  
CC disorder associated with increased expression or activity of NEUAP. NEUAP  
CC or their fragments or derivatives are useful for treating neurological  
CC disorder such as epilepsy, ischemic cerebrovascular disease, stroke,  
CC cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's  
CC disease, dementia and Parkinson's disease. NEUAPs are also useful for  
CC treating other demyelinating diseases, bacterial and viral meningitis,  
CC prion diseases including kuru, Creutzfeldt-Jakob disease, nutritional and  
CC metabolic diseases of the nervous system, neurofibromatosis, other  
CC developmental disorders of the central nervous system, cerebral palsy,  
CC neuroskeletal disorders, autonomic nervous system disorders, cranial  
CC nerve disorders, spinal cord diseases, muscular dystrophy and other  
CC neuromuscular disorders, peripheral neuropathies, mental disorders, inherited,  
CC metabolic, endocrine, and toxic myopathies, a cell proliferative disorder  
CC mood, anxiety and schizophrenic disorders, a cell proliferative disorder  
CC such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis,  
CC cirrhosis, hepatitis, mixed connective tissue disease (MCTD),  
CC myelofibrosis, paroxysmal nocturnal haemoglobinuria, cancers of the  
CC adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, and an  
CC autoimmune/inflammatory disorder such as acquired immunodeficiency  
CC syndrome (AIDS), Addison's disease, adult respiratory distress syndrome,  
CC allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, Werner  
CC syndrome, complications of cancer, hemodialysis, and extracorporeal  
CC circulation, viral, bacterial, fungal parasitic, protozoal, and  
CC helminthic infections, and trauma. This sequence was given the Incyte ID  
CC no. 1871288CB1  
XX  
SQ Sequence 1061 BP; 225 A; 342 C; 326 G; 168 T; 0 U; 0 Other;

Query Match 100.0%; Score 1061; DB 3; Length 1061;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1061; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCTCCTTGGGTTGAAAGCGCTTGGGGTTCAGTGGGCGCATGATCCCCGAGCTGC 60  
Db 1 TCCTCCTTGGGTTGAAAGCGCTTGGGGTTCAGTGGGCGCATGATCCCCGAGCTGC 60  
QY 61 TGGAGAACTGAAGCGGACGGTCTCTCGGAACACGAGCAATGGCGGAGTGGATTGT 120  
Db 61 TGGAGAACTGAAGCGGACGGTCTCTCGGAACACGAGCAATGGCGGAGTGGATTGT 120  
QY 121 TCAGATCATCATCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 180  
Db 121 TCAGATCATCATCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 180  
QY 181 GAGCCACTACAGCTCTCTGACGGTCTTTCATCAGCGGCGGACAGCCAGCGGAGGAG 240  
Db 181 GAGCCACTACAGCTCTCTGACGGTCTTTCATCAGCGGCGGACAGCCAGCGGAGGAG 240  
QY 241 AGAAGATGCCCTGTCTCAGAAGGATGCTTGGCCCTCGGAGAGCACAGTGTTCAGGCAA 300  
Db 241 AGAAGATGCCCTGTCTCAGAAGGATGCTTGGCCCTCGGAGAGCACAGTGTTCAGGCAA 300  
QY 301 CGGAATCCAGAGCGGAGGCTTACGCGCCCGCTCGGCGCCACCGAGCGGCGGCTGCC 360  
Db 301 CGGAATCCAGAGCGGAGGCTTACGCGCCCGCTCGGCGCCACCGAGCGGCGGCTGCC 360  
QY 361 GCCCTTCGCGCCAGCGGAGCGCTTCCACCGCTTCCAGCGCCACCTATCCGTACCTGCAGCA 420  
Db 361 GCCCTTCGCGCCAGCGGAGCGCTTCCACCGCTTCCAGCGCCACCTATCCGTACCTGCAGCA 420

QY 421 CGAGATCGACCTGCGCGCCACCATCTCGCTGTACAGCGGGAGGAGCCCCACCTACCA 480  
Db 421 CGAGATCGACCTGCGCGCCACCATCTCGCTGTACAGCGGGAGGAGCCCCACCTACCA 480  
QY 481 GGGCCCCCTGCACCTCCAGCTTCCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGAGTC 540  
Db 481 GGGCCCCCTGCACCTCCAGCTTCCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGAGTC 540  
QY 541 GGTGCGCGCACCCCAAAACAGAACCATCTTCGACAGTACCTGATGATAGTCCAGGCT 600  
Db 541 GGTGCGCGCACCCCAAAACAGAACCATCTTCGACAGTACCTGATGATAGTCCAGGCT 600  
QY 601 GGGCGGCCCCCTGCGCCCCCAGCAGTAACCTCGGCGCATCAGCGCCACCTGCTACGCGAGCGG 660  
Db 601 GGGCGGCCCCCTGCGCCCCCAGCAGTAACCTCGGCGCATCAGCGCCACCTGCTACGCGAGCGG 660  
QY 661 CGGGCGCATGAGGGGCGCGCGCCACCTACAGCAGGTCATCGGCGCATCAGCGGCGGTC 720  
Db 661 CGGGCGCATGAGGGGCGCGCGCCACCTACAGCAGGTCATCGGCGCATCAGCGGCGGTC 720  
QY 721 CTCTTCCAGCACAGCAGCAGTGGGCGCGCCCTCTTCTGCTGGAGGGGACCCCGCTCCA 780  
Db 721 CTCTTCCAGCACAGCAGCAGTGGGCGCGCCCTCTTCTGCTGGAGGGGACCCCGCTCCA 780  
QY 781 CCACACACATCGCGCCCCCTTAGAGAGCGGCGCATCTGGAGCAAGAGAGGATAAACA 840  
Db 781 CCACACACATCGCGCCCCCTTAGAGAGCGGCGCATCTGGAGCAAGAGAGGATAAACA 840  
QY 841 GAAAGGACACCTCTCTTAGGGTCCCGAGGGGCGCGGCTGGGCTGGTGGTGGTGGTGGT 900  
Db 841 GAAAGGACACCTCTCTTAGGGTCCCGAGGGGCGCGGCTGGGCTGGTGGTGGTGGTGGT 900  
QY 901 GCAGAACACTCCGCGCTTCTTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
Db 901 GCAGAACACTCCGCGCTTCTTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
QY 961 ATCGTGTGGCCCTCCCTCCACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
Db 961 ATCGTGTGGCCCTCCCTCCACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
QY 1021 TGAATGCACAGCTTAAGAGAGCTTTCGCAAAAAA 1061  
Db 1021 TGAATGCACAGCTTAAGAGAGCTTTCGCAAAAAA 1061

RESULT 2

ADC37324

ID ADC37324 standard; DNA; 1085 BP.

XX ADC37324;

XX 18-DEC-2003 (first entry)

XX Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 157.

XX Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;  
KW cancer; infectious disease; bone disease; AIDS;  
KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;  
KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;  
KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.

OS Homo sapiens.

XX WO2003048202-A2.

PN 12-JUN-2003.

XX 03-DEC-2002; 2002WO-JP012644.

XX 03-DEC-2001; 2001JP-00368692.

PR 05-DEC-2001; 2001US-0335829P.

PR 03-OCT-2002; 2002JP-00291302.

PR 04-OCT-2002; 2002US-0415769P.



XX (ASAH ) ASahi KASEI KK.  
XX PA Matsuda A, Muramatsu S;  
XX PI WPI; 2003-505282/47.  
XX DR P-PSDB; ADC37325.  
XX New purified protein that activates nuclear factor kappa B (NF-kappaB),  
PT useful for treating inflammation, autoimmune diseases, cancers,  
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or  
PT ischemic disorders.  
XX Claim 4; SEQ ID NO 157; 938pp; English.  
XX PS  
XX The present invention relates to novel proteins and their coding  
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-  
CC kappaB). The proteins and their coding sequences are useful for treating  
CC a disease associated with NF-kappaB activation, such as inflammation,  
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,  
CC neurodegenerative diseases, or ischaemic disorders.  
XX SQ Sequence 1085 BP; 223 A; 352 C; 334 G; 176 T; 0 U; 0 Other;  
Query Match 99.6%; Score 1057; DB 9; Length 1085;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1057; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCCTCTTGGGTTGCGGTGAAAGCGCTTGGGGGTTCACTGGGCCCATGATCCCGAGCTGC 60  
DB 28 TCCTCTTGGGTTGCGGTGAAAGCGCTTGGGGGTTCACTGGGCCCATGATCCCGAGCTGC 87  
QY 61 TGGAGAACTGAAGCGGAGCGGTCTCTCTGCGAAACCAAGCAATGGCGAGCTGAGTTTGT 120  
DB 88 TGGAGAACTGAAGCGGAGCGGTCTCTCTGCGAAACCAAGCAATGGCGAGCTGAGTTTGT 147  
QY 121 TCAGATCATCATCATCTGTGGTGTGATGATGCTGTGATGCTGTGATGATGCTGTGCT 180  
DB 148 TCAGATCATCATCATCTGTGGTGTGATGATGCTGTGATGCTGTGATGATGCTGTGCT 207  
QY 181 GAGCCACTACAAGCTGTCTGACGGTCTTTCATCAGCCGGCACAGCCAGGGGGGAGGAG 240  
DB 208 GAGCCACTACAAGCTGTCTGACGGTCTTTCATCAGCCGGCACAGCCAGGGGGGAGGAG 267  
QY 241 AGAAGATGCCCTGTCTCAGAAAGGATGCTGTGGCCCTCGGAGAGCACAGTGTGAGGCAA 300  
DB 268 AGAAGATGCCCTGTCTCAGAAAGGATGCTGTGGCCCTCGGAGAGCACAGTGTGAGGCAA 327  
QY 301 CGGAATCCAGAGCGGAGGCTTACGCGCCCGCTCGGCCACCGACCGCTGGCGGTC 360  
DB 328 CGGAATCCAGAGCGGAGGCTTACGCGCCCGCTCGGCCACCGACCGCTGGCGGTC 387  
QY 361 GCCCTTGGCCAGCGGAGGCTTCCACCGCTTCCAGCCCACTATCCGTACCTGCAGCA 420  
DB 388 GCCCTTGGCCAGCGGAGGCTTCCACCGCTTCCAGCCCACTATCCGTACCTGCAGCA 447  
QY 421 CGAGATCGACCTGCCGCCCACTATCTCGCTGTGACGGGGAGGAGCCCACTACCA 480  
DB 448 CGAGATCGACCTGCCGCCCACTATCTCGCTGTGACGGGGAGGAGCCCACTACCA 507  
QY 481 GGGCCCCCTGACCCCTCCAGCTTCCGGACCCCGAGCAGAGCTGGAACCTGAACCGGGAGTC 540  
DB 508 GGGCCCCCTGACCCCTCCAGCTTCCGGACCCCGAGCAGAGCTGGAACCTGAACCGGGAGTC 567  
QY 541 GGTGCGCGCACCCCAACAGAACCATCTTCGACAGTGAACCTGATGATGATGATGATGATGAT 600  
DB 568 GGTGCGCGCACCCCAACAGAACCATCTTCGACAGTGAACCTGATGATGATGATGATGATGAT 627  
QY 601 GGGGGCCCCCTGCCCCCAGCAGTAACCTCGGGCATCAGGCCACCTGCTACGGCAGCGG 660  
DB 628 GGGGGCCCCCTGCCCCCAGCAGTAACCTCGGGCATCAGGCCACCTGCTACGGCAGCGG 687  
QY 661 CGGGCGCATGAGGGGCGCGCCGCCACCTACAGGAGGTGATCGGCCACTACCCCGGGGTC 720

DB 688 CGGGCGCATGAGGGGCGCGCCACCTACAGGAGGTGATCGCCACTACCCGGGGTC 747  
QY 721 CTCCTTCCAGCACCCAGCAGAGCAGTGGCCCGCCCTCTCTTCTGCTGAGGGGACCCGGCTCCA 780  
DB 748 CTCCTTCCAGCACCCAGCAGAGCAGTGGCCCGCCCTCTCTTCTGCTGAGGGGACCCGGCTCCA 807  
QY 781 CCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACA 840  
DB 808 CCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACA 867  
QY 841 GAAAGGACACCTCTCTAGGTTCCCGAGGGGGCGGGCTGGGCTGCTAGGTGAAAAG 900  
DB 868 GAAAGGACACCTCTCTAGGTTCCCGAGGGGGCGGGCTGGGCTGCTAGGTGAAAAG 927  
QY 901 GCAGAAACACTCCCGCTTCTTAGAAGAGAGTGAGAGGAAGCGGGGGCGCAGCAACGC 960  
DB 928 GCAGAAACACTCCCGCTTCTTAGAAGAGAGTGAGAGGAAGCGGGGGCGCAGCAACGC 987  
QY 961 ATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATATTACATGTGATGTCTGGTC 1020  
DB 988 ATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATATTACATGTGATGTCTGGTC 1047  
QY 1021 TGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 1057  
DB 1048 TGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 1084  
RESULT 3  
ABZ36103  
ID ABZ36103 standard; cDNA; 1334 BP.  
XX  
AC ABZ36103;  
XX  
DT 10-FEB-2003 (first entry)  
XX  
DE Human secretory polynucleotide SPTM SEQ ID NO 267.  
XX  
KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;  
KW asthma; Crohn's disease; neurological disorder; epilepsy; cancer;  
KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;  
KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;  
KW anti-inflammatory; immunosuppressive; neuroprotective; nootropic;  
KW neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;  
KW antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;  
XX secretory polynucleotide; secretory protein; gene; ss.  
OS Homo sapiens.  
XX  
PN WO200283876-A2.  
XX  
PD 24-OCT-2002.  
XX  
XX 27-MAR-2002; 2002WO-US009921.  
PF  
XX  
PR 29-MAR-2001; 2001US-0280067P.  
PR 29-MAR-2001; 2001US-0280068P.  
PR 16-MAY-2001; 2001US-0291280P.  
PR 17-MAY-2001; 2001US-0291829P.  
PR 17-MAY-2001; 2001US-0291849P.  
PR 19-JUN-2001; 2001US-0299428P.  
PR 20-JUN-2001; 2001US-0299776P.  
PR 20-JUN-2001; 2001US-030001P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;  
PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Anshey SR;  
PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;  
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;  
PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;  
XX  
DR WPI; 2003-075543/07.

DR P-PSDB; ABP75660.

XX New human secretory proteins and polynucleotides, useful for diagnosing,

PT treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),

PT neurological disorders (e.g. Alzheimer's), or cell proliferations or

PT cancers.

XX

PS Claim 1; SEQ ID NO 267; 458pp + Sequence Listing; English.

XX

CC The invention relates to a secretory polynucleotide (designated aptm)

CC comprising any of 567 polynucleotide sequences (AB235837-AB236403), a

CC naturally occurring polynucleotide sequence at least 90 % identical to

CC the polynucleotide sequence, a polynucleotide complementary to them or an

CC RNA equivalent of them. The polypeptide or polynucleotide are useful for

CC treating, preventing or diagnosing a disease or condition associated with

CC the expression of functional SPTM. These are particularly useful for

CC diagnosing, treating or preventing autoimmune/inflammatory disorders

CC (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's

CC disease), neurological disorders (e.g. epilepsy, Huntington's disease,

CC dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,

CC multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,

CC schizophrenia or amnesia), or cell proliferative disorders (e.g.

CC psoriasis, polycythemia vera, or cancers including adenocarcinoma,

CC leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,

CC breast, cervix or prostate). Note: The sequence data for this patent did

CC not form part of the printed specification, but was obtained in electronic

CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX

SQ Sequence 1334 BP; 299 A; 381 C; 394 G; 260 T; 0 U; 0 Other;

Query Match 99.5%; Score 1056; DB 7; Length 1334;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1056; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCTCTTGGGTTGGGTTGAAAGCGCTTGGGGTTTCAGTGGCCCATGATCCCGAGCTGC 60

DB 1 TCCTCTTGGGTTGGGTTGAAAGCGCTTGGGGTTTCAGTGGCCCATGATCCCGAGCTGC 60

QY 61 TGGAGAACTGAAGCGGACGGTCTCTCTGCGAAACCCAGGCAATGGCGGAGCTGGAGTTGT 120

DB 61 TGGAGAACTGAAGCGGACGGTCTCTCTGCGAAACCCAGGCAATGGCGGAGCTGGAGTTGT 120

QY 121 TCAGATCATCATCATCTGTTGGTGGTATGATGTTGGTGGTGGTGGTGGTGGTGGTGGT 180

DB 121 TCAGATCATCATCATCTGTTGGTGGTATGATGTTGGTGGTGGTGGTGGTGGTGGTGGT 180

QY 181 GAGCCACTACAAGCTGTCTGACGGTCTCTTCATCAGCGGCACAGCCAGCGGCGAGGAG 240

DB 181 GAGCCACTACAAGCTGTCTGACGGTCTCTTCATCAGCGGCACAGCCAGCGGCGAGGAG 240

QY 241 AGAAGATCCCTGTCTCTCAGAGGATGCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAA 300

DB 241 AGAAGATCCCTGTCTCTCAGAGGATGCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAA 300

QY 301 CGGAATCCAGAGCGCGGCTCTACGCCCCCGCTCGGCCCCACCGCCCTGGCGTGGCC 360

DB 301 CGGAATCCAGAGCGCGGCTCTACGCCCCCGCTCGGCCCCACCGCCCTGGCGTGGCC 360

QY 361 GCGCTTCGCCCCAGCGGAGCGCTTCACCGCTTCAGCGCCACCTATCCGTACCTGACGA 420

DB 361 GCGCTTCGCCCCAGCGGAGCGCTTCACCGCTTCAGCGCCACCTATCCGTACCTGACGA 420

QY 421 CGAGATCGACCTGCGCCGCCACCATCTCGCTGTGACAGCGGGAGAGCCCCACCTACCA 480

DB 421 CGAGATCGACCTGCGCCGCCACCATCTCGCTGTGACAGCGGGAGAGCCCCACCTACCA 480

QY 481 GGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCTGGAACCTGAACCGGGAGTC 540

DB 481 GGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCTGGAACCTGAACCGGGAGTC 540

QY 541 GGTGGCGGCAACCCCAACAGAACCATCTTCGACAGTGAAGTGGATGATGATGATGATGAT 600

DB 541 GGTGGCGGCAACCCCAACAGAACCATCTTCGACAGTGAAGTGGATGATGATGATGATGAT 600

QY 601 GGGCGGCCCCCTGCCCCCAGCAGTAACCTCGGCACTCGGCACTCAGGCTACGGCAGCG 660

DB 601 GGGCGGCCCCCTGCCCCCAGCAGTAACCTCGGCACTCAGGCTACGGCAGCG 660

QY 661 CGGGCGCATGGAGGGCGCGCCACCTACAGCGAGGTCTCGGCACTACCGGGGTC 720

DB 661 CGGGCGCATGGAGGGCGCGCCACCTACAGCGAGGTCTCGGCACTACCGGGGTC 720

QY 721 CTCCTTCCAGCACCGCAGCAGTGGGCGCCCTCTTGGAGGGGACCGGCTCCA 780

DB 721 CTCCTTCCAGCACCGCAGCAGTGGGCGCCCTCTTGGAGGGGACCGGCTCCA 780

QY 781 CCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAACA 840

DB 781 CCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAACA 840

QY 841 GAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGCTGGGCTGAAAG 900

DB 841 GAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGCTGGGCTGAAAG 900

QY 901 GCAGAACACTCCGCGCTTCTTAGAGAGGAGTGTAGAGAGGGCGGCGCAGCAACGC 960

DB 901 GCAGAACACTCCGCGCTTCTTAGAGAGGAGTGTAGAGAGGGCGGCGCAGCAACGC 960

QY 961 ATCGTGTGGCCCTCCCTCCACCTCCCTGTTGTAATAATTTACATGTGTGTC 1020

DB 961 ATCGTGTGGCCCTCCCTCCACCTCCCTGTTGTAATAATTTACATGTGTGTC 1020

QY 1021 TGAATGCACAGCTAAGAGAGCTTGCACAAAAA 1056

DB 1021 TGAATGCACAGCTAAGAGAGCTTGCACAAAAA 1056

RESULT 4

AAA75151

ID AAA75151 standard; cDNA; 969 BP.

AC AAA75151;

DT 15-JAN-2001 (first entry)

XX cDNA encoding a human TANGO 261 polypeptide.

DB TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;

XX cellular proliferation; cellular differentiation; cellular adhesion;

KW von Willebrand factor-associated disorder; cell trafficking; cancer;

KW hematopoietic associated disease; atelectasis; pulmonary congestion;

KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;

KW intestinal disorder; spleen associated disease; renal disorder;

KW cardiovascular disorder; ischemic heart disease; hydrocephalus;

KW brain herniation; iatrogenic disease; inflammation; meningitis;

KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;

KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.

XX Homo sapiens.

OS Key Location/Qualifiers

XX CDS 6..764

FT /\*tag= a

FT /product= "TANGO 261"

FT sig\_peptide 6..89

FT /\*tag= b

FT mat\_peptide 90..764

FT /\*tag= c

XX WO200052022-A1.

PN 08-SEP-2000.

XX 01-MAR-2000; 2000WO-US005226.

XX 01-MAR-1999; 99US-0122458P.

















QY 201 CACGGTCTTTCATCAGCCGGCACAGCCAGGGCGGAGGAGAGAGATGCCCTGTCTCTCAG 260  
Db 196 CACGGTCTTTCATCAGCCGGCACAGCCAGGGCGGAGGAGAGAGATGCCCTGTCTCTCAG 255  
QY 261 AAGGATGCTGTGGCCCTCGGAGAGCAGTGTCTCAGCAACGGAATCCAGAGCGCAGG 320  
Db 256 AAGGATGCTGTGGCCCTCGGAGAGCAGTGTCTCAGCAACGGAATCCAGAGCGCAGG 315  
QY 321 TCTACGCCCGCTCGGCCACCGACCGCTGCGCGCTGCGCGCTGCGCGCGGAGC 380  
Db 316 TCTACGCCCGCTCGGCCACCGACCGCTGCGCGCTGCGCGCTGCGCGCGGAGC 375  
QY 381 GCTTCCACCGCTTCCAGCCACCTATCCGTAACCTGACGACGAGATCGACCTGCGGCCA 440  
Db 376 GCTTCCACCGCTTCCAGCCACCTATCCGTAACCTGACGACGAGATCGACCTGCGGCCA 435  
QY 441 CCATCTCGTGTGAGAGGGGAGGAGCCCGACCTACCGGGCCCTGACCCCTCCAGC 500  
Db 436 CCATCTCGTGTGAGAGGGGAGGAGCCCGACCTACCGGGCCCTGACCCCTCCAGC 495  
QY 501 TTCCGGACCCGAGCAGCTGGAACCTGACCGGGAGTGGTGGCGGACCCCAACA 560  
Db 496 TTCCGGACCCGAGCAGCTGGAACCTGACCGGGAGTGGTGGCGGACCCCAACA 555  
QY 561 GAACCATCTTCGACAGTGACCTGATGAGATAGTGGCGGCTGCGGCCCTGCCCCCA 620  
Db 556 GAACCATCTTCGACAGTGACCTGATGAGATAGTGGCGGCTGCGGCCCTGCCCCCA 615  
QY 621 GCAGTAACCTCGGGCATCAGCGCCACCTGCTACGGCAGCGCGGGGCGATGAGGGCGC 680  
Db 616 GCAGTAACCTCGGGCATCAGCGCCACCTGCTACGGCAGCGCGGGGCGATGAGGGCGC 675  
QY 681 CGCCACCTACAGCGAGGTATCGGCCATACCGGGGTCTCTCTTCAGCACCAGCAGA 740  
Db 676 CGCCACCTACAGCGAGGTATCGGCCATACCGGGGTCTCTCTTCAGCACCAGCAGA 735  
QY 741 GCAGTGGCGCGCCCTCTCTGTTGGAGGGGACCGGCTCCACACACACATCGCGCCC 800  
Db 736 GCAGTGGCGCGCCCTCTCTGTTGGAGGGGACCGGCTCCACACACACATCGCGCCC 795  
QY 801 TAGAGAGCGCAGCCATCTGAGCAAGAGAGATAAACAGAAAGGACACCTCTCTTAGG 860  
Db 796 TAGAGAGCGCAGCCATCTGAGCAAGAGAGATAAACAGAAAGGACACCTCTCTTAGG 855  
QY 861 GTCCCCAGGGGGCGGGCTGGGGCTGCTAGTGAAAGGAGCAGACACTCGCGCTTCT 920  
Db 856 GTCCCCAGGGGGCGGGCTGGGGCTGCTAGTGAAAGGAGCAGACACTCGCGCTTCT 915  
QY 921 TAGAAGAGGAGTGAGAGGAGGGCGGGGCGCAGCAACGATCGTGTGGCCCTCCCTCC 980  
Db 916 TAGAAGAGGAGTGAGAGGAGGGCGGGGCGCAGCAACGATCGTGTGGCCCTCCCTCC 975  
QY 981 CACCTCCCTGTGTATAAATATTACATGTGATGTCTGGTCTGAATGCACAGCTAAGAGA 1040  
Db 976 CACCTCCCTGTGTATAAATATTACATGTGATGTCTGGTCTGAATGCACAGCTAAGAGA 1035  
QY 1041 GCTTGCAAAAAA 1057  
Db 1036 GCTTGCAAAAAA 1052

RESULT 9  
ACC49536  
ID ACC49536 standard; cDNA; 1850 BP.  
XX AC ACC49536;  
XX AC  
DT 01-JUL-2003 (first entry)  
XX  
DE Tumour-associated antigenic target protein TAT180 cDNA SEQ ID NO:44.  
XX  
KW Human; tumour-associated antigenic target; TAT; tumour; diagnosis;  
KW cancer; gene; ss.

XX Homo sapiens.  
OS WO2003024392-A2.  
XX  
PN 27-MAR-2003.  
XX  
PD 11-SEP-2002; 2002WO-US028859.  
XX  
PR 18-SEP-2001; 2001US-0323268P.  
PR 19-OCT-2001; 2001US-0339227P.  
PR 07-NOV-2001; 2001US-0336827P.  
PR 20-NOV-2001; 2001US-0331906P.  
PR 02-JAN-2002; 2002US-0345444P.  
PR 03-APR-2002; 2002US-0369724P.  
PR 19-AUG-2002; 2002US-0404809P.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Frantz G, Hillan KJ, Phillips HS, Polakis P, Spencer SD;  
PI Williams PM, Wu TD, Zhang Z;  
XX  
DR WPI; 2003-354551/33.  
DR P-PSDB; ABP97218.  
XX  
PT New antibodies against tumor-associated antigenic target polypeptide,  
PT useful for treating or diagnosing tumors or cancers in mammals, e.g.  
PT prostate cancer, lung cancer, prostate adenocarcinomas or renal cell  
PT carcinomas.  
XX  
PS Claim 2; Fig 44; 285pp; English.  
XX  
CC ACC49493 to ACC49552 encode the human tumour-associated antigenic target  
CC (TAT) proteins given in ABP97175 to ABP97234. The present invention  
CC describes an isolated antibody that binds to a polypeptide having at  
CC least 80 % sequence identity to any of the 60 150-800 residue amino acid  
CC sequences (S1), given in ABP97175 to ABP97234, comprising (S1), lacking  
CC its associated signal peptide, encoded by any of the 60 2000-3000 base  
CC pair sequences (S2), given in ACC49493 to ACC49552. TAT proteins have  
CC cytostatic activity. The antibody can be used for treating or diagnosing  
CC tumours or cancers in mammals; e.g. prostate cancer, lung cancer, breast  
CC cancer, colon cancer, ovarian cancer, prostate adenocarcinomas, renal  
CC cell carcinomas, or thyroid cancer  
XX  
SQ Sequence 1850 BP; 477 A; 472 C; 498 G; 403 T; 0 U; 0 Other;  
Query Match 87.3%; Score 926; DB 7; Length 1850;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 976; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 81 GTCTCTCTCGGAAACAGGCAATGGCGGAGCTGAGTTGTTTCAGATCATCATCGTGG 140  
Db 76 GTCTCTCTCGGAAACAGGCAATGGCGGAGCTGAGTTGTTTCAGATCATCATCGTGG 135  
QY 141 TGGTGATGATGGTGATGGTGATGATGATGATGATGATGATGATGATGATGATGATG 200  
Db 136 TGGTGATGATGGTGATGGTGATGATGATGATGATGATGATGATGATGATGATGATG 195  
QY 201 CACGGTCTTTCATCAGCCGGCACAGCCAGGGCGGAGGAGAGAGATGCCCTGTCTCTCAG 260  
Db 196 CACGGTCTTTCATCAGCCGGCACAGCCAGGGCGGAGGAGAGAGATGCCCTGTCTCTCAG 255  
QY 261 AAGGATGCTGTGGCCCTCGGAGAGCAGTGTCTCAGCAACGGAATCCAGAGCGCAGG 320  
Db 256 AAGGATGCTGTGGCCCTCGGAGAGCAGTGTCTCAGCAACGGAATCCAGAGCGCAGG 315  
QY 321 TCTACGCCCGCTCGGCCACCGACCGCTGCGCGCTGCGCGCTGCGCGCGGAGC 380  
Db 316 TCTACGCCCGCTCGGCCACCGACCGCTGCGCGCTGCGCGCTGCGCGCGGAGC 375  
QY 381 GCTTCCACCGCTTCCAGCCACCTATCCGTAACCTGACGACGAGATCGACCTGCGGCCA 440  
Db 376 GCTTCCACCGCTTCCAGCCACCTATCCGTAACCTGACGACGAGATCGACCTGCGGCCA 435

QY	441	CCATCTCGCTGT	CAGACGGGAGGAGCCCCCACCCTTACAGGGCCCCTGCACCCCTCCAGC	500
DB	436	CCATCTCGCTGT	CAGACGGGAGGAGCCCCCACCCTTACAGGGCCCCTGCACCCCTCCAGC	495
QY	501	TTCGGGACCCC	GAGCAGACTGGAACTGAAACCGGGAGTCGGTGCGCGCACCCCCCAACA	560
DB	496	TTCGGGACCCC	GAGCAGACTGGAACTGAAACCGGGAGTCGGTGCGCGCACCCCCCAACA	555
QY	561	GAACCATCTTC	GACAGTGAATGGATAGTGCACGGCTGGGCGGCCCTTGCCCCCCCCA	620
DB	556	GAACCATCTTC	GACAGTGAATGGATAGTGCACGGCTGGGCGGCCCTTGCCCCCCCCA	615
QY	621	GCAGTAAC	TGCGGCAATCAGGCCACGTGCTACGGCAGCGCGGGCGGCATGGAGGGGCGCG	680
DB	616	GCAGTAAC	TGCGGCAATCAGGCCACGTGCTACGGCAGCGCGGGCGGCATGGAGGGGCGCG	675
QY	681	CGCCACCTAC	AGCGAGGTCAATCGGCCACTACCCGGSGTCTCTCTTCCAGCACCAAGAGA	740
DB	676	CGCCACCTAC	AGCGAGGTCAATCGGCCACTACCCGGSGTCTCTCTTCCAGCACCAAGAGA	735
QY	741	GCAGTGGGCG	CCCTCTCTGCTGGAGGGGACCCGGGTCCACACACACATCGGCGCC	800
DB	736	GCAGTGGGCG	CCCTCTCTGCTGGAGGGGACCCGGGTCCACACACACATCGGCGCC	795
QY	801	TAGAGAGCGC	AGCCATCTGGAGCAAAGAGATAAAACAGAAAGAACACCTCTCTTAGG	860
DB	796	TAGAGAGCGC	AGCCATCTGGAGCAAAGAGATAAAACAGAAAGAACACCTCTCTTAGG	855
QY	861	GTCCCCAGGG	GGGCGCGGCTGGGGCTGCGTAGGTGAAAAAGGCAGAACACTCCGGGCTTCT	920
DB	856	GTCCCCAGGG	GGGCGCGGCTGGGGCTGCGTAGGTGAAAAAGGCAGAACACTCCGGGCTTCT	915
QY	921	TAGAAAGAGG	AGTGAGAGNAAGGCGGGGGCGCAGCAACCCATCGTGTGGCCCTCCCTCC	980
DB	916	TAGAAAGAGG	AGTGAGAGNAAGGCGGGGGCGCAGCAACCCATCGTGTGGCCCTCCCTCC	975
QY	981	CACCTCCCTG	TGTATAAATATTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGA	1040
DB	976	CACCTCCCTG	TGTATAAATATTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGA	1035
QY	1041	GCTTGCAAAA	AAAAAAAA 1057	
DB	1036	GCTTGCAAAA	AAAAAAAA 1052	

**RESULT 10**

AAD60105  
ID AAD60105 standard; cDNA; 1140 BP.

US6566130-B1.

XX  
PD  
20-MAY-2003

XX  
PF 26-JAN-2001; 2001US-00769482.



QY 755 TCCTTGCTGGAGGACCCGGCTCCACCACACACATCGCGCCCTAGAGAGCGCAGCC 814  
Db |||||  
QY 749 TCCTTGCTGGAGGACCCGGCTCCACCACACACATCGCGCCCTAGAGAGCGCAGCC 808  
Db |||||  
QY 815 ATCTGGAGCAAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCAGGGGGC 874  
Db |||||  
QY 809 ATCTGGAGCAAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCAGGGGGC 868  
Db |||||  
QY 875 CGGGCTGGGCTGCTAGGTGAAAAGGCGAGCACTCGCGCTTCTTAGAGAGGAGTGA 934  
Db |||||  
QY 869 CGGGCTGGGCTGCTAGGTGAAAAGGCGAGCACTCGCGCTTCTTAGAGAGGAGTGA 928  
Db |||||  
QY 935 GAGGAAGCGGGGGCGCAGCAACGCAATCGTGTGGCCCTCCCTCCACCTCCCTGTGTA 994  
Db |||||  
QY 929 GAGGAAGCGGGGGCGCAGCAACGCAATCGTGTGGCCCTCCCTCCACCTCCCTGTGTA 988  
Db |||||  
QY 995 TAAATATTTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAAA 1054  
Db |||||  
QY 989 TAAATATTTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAAA 1048  
Db |||||  
QY 1055 AAA 1057  
Db |||||  
QY 1049 AAA 1051

RESULT 11  
AAA75164  
ID AAA75164 standard; cDNA; 969 BP.  
AC AAA75164;  
XX  
DT 15-JAN-2001 (first entry)  
XX  
DE cDNA clone encoding a human TANGO 261 polypeptide.  
XX  
KW TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;  
KW cellular proliferation; cellular differentiation; cellular adhesion;  
KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
KW intestinal disorder; spleen associated disease; renal disorder;  
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
KW brain herniation; iatrogenic disease; inflammation; meningitis;  
KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;  
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.  
OS Homo sapiens.  
XX  
Key Location/Qualifiers  
CDS 6..764  
FT /\*tag= a  
FT /product= "TANGO 261"  
XX  
WO200052022-A1.  
PN  
XX  
PD 08-SEP-2000.  
XX  
PF 01-MAR-2000; 2000WO-US005226.  
XX  
PR 01-MAR-1999; 99US-0122458P.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;  
XX  
XX WPI; 2000-579269/54.  
DR P-PSDB; AAB18462.  
XX  
PT Novel human and murine secreted proteins designated TANGO 216, 261, 262,  
PT 266 and 267 useful as modulating agents of cellular processes, e.g. for  
PT treating cancer.  
XX  
PS Disclosure; Page; 175pp; English.

XX  
CC AAA75163-65 encode human TANGO 261 proteins. The specification also  
CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO  
CC polypeptides can be used to modulate cellular proliferation, modulate  
CC cellular differentiation and/or modulate cellular adhesion. The proteins  
CC can be used to treat any von Willebrand factor-associated disorder,  
CC regulate extracellular matrix structuring, cellular adhesion, and cell  
CC trafficking and/or migration, modulate cellular interactions, modulate  
CC cell adhesion in proliferative disorders, such as cancer, modulate the  
CC proliferation, differentiation, and/or function of cells that appear in  
CC the bone marrow, and leukocytes, treat bone marrow, blood and  
CC hematopoietic associated diseases and disorders, atelectasis, pulmonary  
CC congestion or oedema, emphysema, chronic bronchitis, bronchial asthma and  
CC bronchiectasis, intestinal disorders, spleen associated diseases,  
CC modulate renal disorders, treat cardiovascular disorders such as ischemic  
CC heart disease, modulate the proliferation, differentiation, and/or  
CC function of bone and cartilage cells and to treat bone and/or cartilage  
CC associated diseases or disorder. They may also be used to treat disorders  
CC associated with the ovaries, and cerebral oedema, hydrocephalus, brain  
CC herniations, iatrogenic disease, inflammations, bacterial and viral  
CC meningitis, Alzheimer's Disease, brain cancers, hydrocephalus and  
CC disease, multiple sclerosis, Parkinson's disease, toxoplasmosis and  
CC encephalitis, and treat hepatic disorders. note: the present sequence  
CC does not appear in the specification; it was created using information  
CC provided  
XX  
SQ Sequence 969 BP; 210 A; 317 C; 294 G; 148 T; 0 U; 0 Other;  
Query Match 85.8%; Score 910; DB 3; Length 969;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 960; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 101 ATGGCGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGTGTGATGATGATGATG 160  
Db |||||  
QY 161 GTGTGATCAGTGCCTGCTGAGCCACTACAGCTGTCTGCAAGTCTCTTCATCAGCGG 220  
Db |||||  
QY 66 GTGTGATCAGTGCCTGCTGAGCCACTACAGCTGTCTGCAAGTCTCTTCATCAGCGG 125  
Db |||||  
QY 221 CACAGCCAGGGCGGAGGAGAGAGATGCCCTGTCTCTCAGAGGATGCTGTGGCCCTCG 280  
Db |||||  
QY 126 CACAGCCAGGGCGGAGGAGAGAGATGCCCTGTCTCTCAGAGGATGCTGTGGCCCTCG 185  
Db |||||  
QY 281 GAGAGCACAGTGTTCAGGCAACGGAATCCAGAGCGCAGGTCTACGCCGCCCTCGGCC 340  
Db |||||  
QY 186 GAGAGCACAGTGTTCAGGCAACGGAATCCAGAGCGCAGGTCTACGCCGCCCTCGGCC 245  
QY 341 ACCGACCGCTGGCCGTGCGGCCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCAGGCC 400  
Db |||||  
QY 246 ACCGACCGCTGGCCGTGCGGCCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCAGGCC 305  
QY 401 ACCTATCCGTACCTGCAGCAGAGATCGACTGCGGCCGCCACCATCTCGTGTTCAGACGG 460  
Db |||||  
QY 306 ACCTATCCGTACCTGCAGCAGAGATCGACTGCGGCCGCCACCATCTCGTGTTCAGACGG 365  
QY 461 GAGGAGCCCCACCTTACCAGGGCCCCCTGCAGCTTCCAGCTTCCAGCGCAGCAG 520  
Db |||||  
QY 366 GAGGAGCCCCACCTTACCAGGGCCCCCTGCAGCTTCCAGCTTCCAGCGCAGCAG 425  
QY 521 CTGGAACCTGAACCGGGAGTGGTGGCGGCCCTTCAGACCCCAACAGAACCATCTTCGACAGT 580  
Db |||||  
QY 426 CTGGAACCTGAACCGGGAGTGGTGGCGGCCCTTCAGACCCCAACAGAACCATCTTCGACAGT 485  
QY 581 CTGATGGATAGTCCAGGCTGGCGGCCCTTCAGACCCCAACAGAACCATCTTCGAGCATCAGC 640  
Db |||||  
QY 486 CTGATGGATAGTCCAGGCTGGCGGCCCTTCAGACCCCAACAGAACCATCTTCGAGCATCAGC 545  
QY 641 GCCACGTGCTACGGCAGCGGGCGGCGCATGGAGGGCGCGGCCCAACCTACAGAGGAGTTC 700  
Db |||||  
QY 546 GCCACGTGCTACGGCAGCGGGCGGCGCATGGAGGGCGCGGCCCAACCTACAGAGGAGTTC 605  
QY 701 ATCGGCCCACTACCGGGGTCTCTTCAGACCCAGCAGAGCAGTGGGCCCTCTCTTG 760





Db 546 GCCACGTGCTACGGCAGCGGGGGCGATGGAGGGGGCGCCGCCACCTACAGCGAGGTC 605  
QY 701 ATCGGCCACTACCGGGGTCTCTTCCAGCACAGCAGAGCAGTGGCGCCCTCTCTTG 760  
Db 606 ATCGGCCACTACCGGGGTCTCTTCCAGCACAGCAGAGCAGTGGCGCCCTCTCTTG 665  
QY 761 CTGGAGGGACCGGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGG 820  
Db 666 CTGGAGGGACCGGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGG 725  
QY 821 AGCAAGAGAGGATATAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGCGGCT 880  
Db 726 AGCAAGAGAGGATATAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGCGGCT 785  
QY 881 GGGGCTGCTAGGTGAAAGGAGCAGACACTCCGGCTTCTTAGAAGAGGAGTGAGAGGA 940  
Db 786 GGGGCTGCTAGGTGAAAGGAGCAGACACTCCGGCTTCTTAGAAGAGGAGTGAGAGGA 845  
QY 941 GGGGGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCTGTGTATAATA 1000  
Db 846 GGGGGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCTGTGTATAATA 905  
QY 1001 TTTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 1060  
Db 906 TTTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 965  
QY 1061 A 1061  
Db 966 A 966

RESULT 13  
ACC49552  
ID ACC49552 standard; cDNA; 4839 BP.  
XX  
AC ACC49552;  
XX  
DT 01-JUL-2003 (first entry)  
XX  
DE Tumour-associated antigenic target protein TAT379 SEQ ID NO:119.  
XX  
KW Human; tumour-associated antigenic target; TAT; tumour; diagnosis;  
KW cancer; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2003024392-A2.  
XX  
PD 27-MAR-2003.  
XX  
PF 11-SEP-2002; 2002WO-US028859.  
XX  
PR 18-SEP-2001; 2001US-0323268P.  
PR 19-OCT-2001; 2001US-0339227P.  
PR 07-NOV-2001; 2001US-0336827P.  
PR 20-NOV-2001; 2001US-0331906P.  
PR 02-JAN-2002; 2002US-0345444P.  
PR 03-APR-2002; 2002US-0369724P.  
PR 19-AUG-2002; 2002US-0404809P.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Frantz G, Hillan KJ, Phillips HS, Polakis P, Spencer SD;  
PI Williams PM, Wu TD, Zhang Z;  
XX  
DR WPI; 2003-354551/33.  
DR P-PSDB; ABP97234.  
XX

PT New antibodies against tumor-associated antigenic target polypeptide,  
PT useful for treating or diagnosing tumors or cancers in mammals, e.g.  
PT prostate cancer, lung cancer, prostate adenocarcinomas or renal cell  
PT carcinomas.  
XX

PS Claim 2; Fig 119; 285pp; English.  
XX  
CC ACC49493 to ACC49552 encode the human tumour-associated antigenic target  
(TAT) proteins given in ABP97175 to ABP97234. The present invention  
CC describes an isolated antibody that binds to a polypeptide having at  
CC least 80 % sequence identity to any of the 60 150-800 residue amino acid  
CC sequences (S1), given in ABP97175 to ABP97234, comprising (S1), lacking  
CC its associated signal peptide, encoded by any of the 60 2000-3000 base  
CC pair sequences (S2), given in ACC49493 to ACC49552. TAT proteins have  
CC cytostatic activity. The antibody can be used for treating or diagnosing  
CC tumours or cancers in mammals, e.g. prostate cancer, lung cancer, breast  
CC cancer, colon cancer, ovarian cancer, prostate adenocarcinomas, renal  
CC cell carcinomas, or thyroid cancer  
XX  
SQ Sequence 4839 BP; 1103 A; 1313 C; 1290 G; 1133 T; 0 U; 0 Other;  
Query Match 85.0%; Score 902; DB 7; Length 4839;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 952; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 105 CGGAGCTGGAGTTGTTTCAGATCATCATCTGTTGGTGTGATGTTGGTGGTGG 164  
Db 430 CGGAGCTGGAGTTGTTTCAGATCATCATCTGTTGGTGTGATGTTGGTGGTGG 489  
QY 165 TGATCACGTGCTGCTGAGCCACTACAAGCTGTCTGACGGTCTTTCATCAGCCGGCACA 224  
Db 490 TGATCACGTGCTGCTGAGCCACTACAAGCTGTCTGACGGTCTTTCATCAGCCGGCACA 549  
QY 225 GCCAGGGCGGAGGAGAGAAGATGCCCTGTCTCCTCAGAAAGGATGCCCTGTCCTCGGAGA 284  
Db 550 GCCAGGGCGGAGGAGAGAAGATGCCCTGTCTCCTCAGAAAGGATGCCCTGTCCTCGGAGA 609  
QY 285 GCACAGTGTGAGGCAACGGAAATCCAGAGCCGCGAGGTCTAGCCGCCCTCGGCCACCG 344  
Db 610 GCACAGTGTGAGGCAACGGAAATCCAGAGCCGCGAGGTCTAGCCGCCCTCGGCCACCG 669  
QY 345 ACCGCCCTGGCGTGCCTTCCAGAGCCGCGAGGTCTTCCAGCCCTTCCAGCCACCT 404  
Db 670 ACCGCCCTGGCGTGCCTTCCAGAGCCGCGAGGTCTTCCAGCCCTTCCAGCCACCT 729  
QY 405 ATCCGTACTGTCAGCAGCAGATCGACCTGCGGCCACCATCTCGCTGTCTCAGACGGGAGG 464  
Db 730 ATCCGTACTGTCAGCAGCAGATCGACCTGCGGCCACCATCTCGCTGTCTCAGACGGGAGG 789  
QY 465 AGCCCCCAGCCTACAGGGCCCTGACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGG 524  
Db 790 AGCCCCCAGCCTACAGGGCCCTGACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGG 849  
QY 525 AACTGAACCGGGAGTGGTGGCGCACCCGCCCAACAGAACCATCTTCGACAGTACCTGA 584  
Db 850 AACTGAACCGGGAGTGGTGGCGCACCCGCCCAACAGAACCATCTTCGACAGTACCTGA 909  
QY 585 TGGATAGTCCAGGCTGGCGGCCCTTGCCCCCAGCAGTAACTCGGGCATCAGCGCCA 644  
Db 910 TGGATAGTCCAGGCTGGCGGCCCTTGCCCCCAGCAGTAACTCGGGCATCAGCGCCA 969  
QY 645 CGTCTACGGCAGCGCGGGCGCATGAGGGGCGCGGCCCACTACAGCGAGGTCTATCG 704  
Db 970 CGTCTACGGCAGCGCGGGCGCATGAGGGGCGCGGCCCACTACAGCGAGGTCTATCG 1029  
QY 705 GCCACTACCGGGGTCTCTCTTCCAGCACAGCAGCAGTGGGCGCCCTCTTCTGCTGG 764  
Db 1030 GCCACTACCGGGGTCTCTCTTCCAGCACAGCAGCAGCAGTGGGCGCCCTCTTCTGCTGG 1089  
QY 765 AGGGGACCGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCA 824  
Db 1090 AGGGGACCGGCTCCACCACACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCA 1149  
QY 825 AAGAGAGGATATAACAGAAAGGACACCTCTCTTAGGGTCCCGAGGGGGCGGGCTGGGG 884  
Db 1150 AAGAGAGGATATAACAGAAAGGACACCTCTCTTAGGGTCCCGAGGGGGCGGGCTGGGG 1209  
QY 885 CTGCGTAGGTGAAAGGACAGACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAGCGG 944















```

; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-769-482-5

```

```

Query Match      2.3%; Score 24; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      900 GGCAGAACACTCCGCGCTTCTTAG 923
Db      1 GGCAGAACACTCCGCGCTTCTTAG 24

```

```

RESULT 4
US-09-769-482-6/c
; Sequence 6, Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-769-482-6

```

```

Query Match      2.3%; Score 24; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1022 GAATGCACAAGCTAAGAGAGCTTG 1045
Db      24 GAATGCACAAGCTAAGAGAGCTTG 1

```

```

RESULT 5
US-09-769-482-10/c
; Sequence 10, Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED

```

```

; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-769-482-10

```

```

Query Match      2.3%; Score 24; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1022 GAATGCACAAGCTAAGAGAGCTTG 1045
Db      24 GAATGCACAAGCTAAGAGAGCTTG 1

```

```

RESULT 6
US-08-332-766A-1
; Sequence 1, Application US/08332766A
; Patent No. 5843647
; GENERAL INFORMATION:
; APPLICANT: JEFFREYS, Alec J.
; APPLICANT: ARMOUR, John
; TITLE OF INVENTION: SIMPLE TANDEM REPEATS
; NUMBER OF SEQUENCES: 125
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,766A
; FILING DATE: 01-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9326052.9
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BIRD, Donald J.
; REGISTRATION NUMBER: 25,323
; REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-332-766A-1

```

```

Query Match      2.2%; Score 23; DB 2; Length 377;

```

Best Local Similarity 100.0%; Pred. No. 2.5; Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 TGGTGGTGATGATGGTGATGGTG 160  
 Db 123 TGGTGGTGATGATGGTGATGGTG 145

RESULT 7

US-09-769-482-7  
 ; Sequence 7, Application US/09769482  
 ; Patent No. 6566130  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SRIVASTAVA, SHIV  
 ; APPLICANT: MOUL, JUDD W.  
 ; APPLICANT: XU, LINDA L.  
 ; APPLICANT: SEGAWA, TAKEHIKO  
 ; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED  
 ; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY  
 ; FILE REFERENCE: 04995.0057-00000  
 ; CURRENT APPLICATION NUMBER: US/09/769,482  
 ; CURRENT FILING DATE: 2001-01-26  
 ; PRIOR APPLICATION NUMBER: 60/178,772  
 ; PRIOR FILING DATE: 2000-01-28  
 ; PRIOR APPLICATION NUMBER: 60/179,045  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 67  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 7  
 ; LENGTH: 22  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Primer  
 US-09-769-482-7

Query Match 2.0%; Score 21; DB 4; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CTTGGGTTCCGGTGAAGCGC 26  
 Db 1 CTTGGGTTCCGGTGAAGCGC 21

RESULT 8

US-08-522-421-7  
 ; Sequence 7, Application US/08522421  
 ; Patent No. 5908973  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Abu-Bakar, Umi Kalsom  
 ; APPLICANT: Barton, Sarah Louise  
 ; APPLICANT: Gallego-Veigas, Pedro Pablo  
 ; APPLICANT: Gray, Julie Elizabeth  
 ; APPLICANT: Grierson, Donald  
 ; APPLICANT: Lowe, Alexandra Louise  
 ; APPLICANT: Picton, Steve  
 ; APPLICANT: Whotton, Lee Colin  
 ; TITLE OF INVENTION: DNA, DNA CONSTRUCTS, CELLS AND PLANTS  
 ; TITLE OF INVENTION: DERIVED THEREFROM  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.  
 ; STREET: 1100 New York Avenue, N.W.  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20005-3918  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Microsoft Word

; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/522,421  
 ; FILING DATE: 11-JAN-1996  
 ; CLASSIFICATION: 800  
 ; PRIOR APPLICATION DATA: GB 9305868.3  
 ; FILING DATE: 22-MAR-1993  
 ; PRIOR APPLICATION DATA: GB 9305869.1  
 ; FILING DATE: 22-MAR-1993  
 ; PRIOR APPLICATION DATA: GB 9305859.2  
 ; FILING DATE: 22-MAR-1993  
 ; PRIOR APPLICATION DATA: GB 9305865.9  
 ; FILING DATE: 22-MAR-1993  
 ; PRIOR APPLICATION DATA: GB 9305866.7  
 ; FILING DATE: 22-MAR-1993  
 ; PRIOR APPLICATION DATA: GB 9305867.5  
 ; FILING DATE: 22-MAR-1993  
 ; PRIOR APPLICATION DATA: GB 9305860.0  
 ; FILING DATE: 22-MAR-1993  
 ; PRIOR APPLICATION DATA: GB 9305862.6  
 ; FILING DATE: 22-MAR-1993  
 ; PRIOR APPLICATION DATA: GB 9314351.9  
 ; FILING DATE: 12-JUL-1993  
 ; PRIOR APPLICATION DATA: GB 9320988.0  
 ; FILING DATE: 12-OCT-1993  
 ; INFORMATION FOR SEQ ID NO: 7:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 686 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: ERT17  
 US-08-522-421-7

Query Match 2.0%; Score 21; DB 2; Length 686;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1041 GCTTGCAAAAAAAAAAAAAA 1061  
 Db 663 GCTTGCAAAAAAAAAAAAAA 683

RESULT 9

US-08-165-315D-3  
 ; Sequence 3, Application US/08165315D  
 ; Patent No. 5525716  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Odd-Arne Olsen  
 ; APPLICANT: Roger Kalla  
 ; TITLE OF INVENTION: Promoter  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: St. Onge, Steward, Johnston & Reens  
 ; STREET: 986 Bedford Street  
 ; CITY: Stamford  
 ; STATE: Connecticut  
 ; COUNTRY: U.S.A.  
 ; ZIP: 06905  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" 1.44 Mb diskette  
 ; COMPUTER: IBM PC

```

; OPERATING SYSTEM: MS DOS
; SOFTWARE: Word Processor
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/165,315D
; FILING DATE: 10 December 1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9324707.0
; FILING DATE: 2 December 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary M. Krinsky
; REGISTRATION NUMBER: 32423
; REFERENCE/DOCKET NUMBER: 2105-P0001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-324-6155
; TELEFAX: 201-327-1096
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1327
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: DNA
; FRAGMENT TYPE: gene
; FEATURE:
; NAME/KEY: Ltp2 gene
US-08-165-315D-3

```

Query Match 2.0%; Score 21; DB 1; Length 1327;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 143 GTGATGATGGTGGTGGTG 163
Db 918 GTGATGATGGTGGTGGTG 938

```

```

RESULT 10
US-08-046-585-15/c
; Sequence 15, Application US/08046585
; Patent No. 5453362
; GENERAL INFORMATION:
; APPLICANT: Lamarco, Kelly
; APPLICANT: Wilson, Angus
; APPLICANT: Herr, Winship
; TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN:
; TITLE OF INVENTION: HOST CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/046,585
; FILING DATE: 12-APR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-57503-1/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299

```

```

; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-046-585-15

```

Query Match 2.0%; Score 21; DB 1; Length 8252;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 138 TGGTGGTGGTGGTGGTGG 158
Db 2748 TGGTGGTGGTGGTGGTGG 2728

```

```

RESULT 11
US-08-393-703-15/c
; Sequence 15, Application US/08393703
; Patent No. 5585239
; GENERAL INFORMATION:
; APPLICANT: Lamarco, Kelly
; APPLICANT: Wilson, Angus
; APPLICANT: Herr, Winship
; TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN:
; TITLE OF INVENTION: HOST CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,703
; FILING DATE: 24-FEB-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-57503-2/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-393-703-15

```

Query Match 2.0%; Score 21; DB 1; Length 8252;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 138 TGGTGGTGGTGGTGGTGG 158
Db 2748 TGGTGGTGGTGGTGGTGG 2728

```

RESULT 12  
 PCT-US93-11721-15/c



Sequence 15, Application PC/TUS93111721  
 GENERAL INFORMATION:  
 APPLICANT: Lamarc, Kelly  
 APPLICANT: Wilson, Angus  
 APPLICANT: Herr, Winship  
 TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN:  
 TITLE OF INVENTION: HOST CELL FACTOR  
 NUMBER OF SEQUENCES: 15  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
 STREET: 4 Embarcadero Center, Suite 3400  
 CITY: San Francisco  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94111-4187  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US93/11721  
 FILING DATE: 03-DEC-1993  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Osman, Richard A  
 REGISTRATION NUMBER: 36,627  
 REFERENCE/DOCKET NUMBER: FP-57503-1/RAO  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 781-1989  
 TELEFAX: (415) 398-3249  
 TELEX: 910 277299  
 INFORMATION FOR SEQ ID NO: 15:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 8252 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 PCT-US93-11721-15

Query Match 2.0%; Score 21; DB 5; Length 8252;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 TGGTGGTGATGATGGTGG 158  
 Db 2748 TGGTGGTGATGATGGTGG 2728

RESULT 13  
 US-08-916-421B-1/c  
 Sequence 1, Application US/08916421B  
 Patent No. 6503729  
 GENERAL INFORMATION:  
 APPLICANT: Bult et al.  
 TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus jannaschii  
 TITLE OF INVENTION: jannaschii  
 FILE REFERENCE: PB275  
 CURRENT APPLICATION NUMBER: US/08/916,421B  
 CURRENT FILING DATE: 1997-08-22  
 PRIOR APPLICATION NUMBER: US 60/024,428  
 PRIOR FILING DATE: 1996-08-22  
 NUMBER OF SEQ ID NOS: 3  
 SOFTWARE: Patentin version 3.1  
 SEQ ID NO 1  
 LENGTH: 1664976  
 TYPE: DNA  
 ORGANISM: Methanococcus jannaschii  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (28222)..(28222)

OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (28257)..(28258)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (84773)..(84773)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (84808)..(84808)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (84812)..(84812)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (98120)..(98120)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (98159)..(98159)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (98239)..(98239)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (98266)..(98266)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (98343)..(98343)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (103998)..(103998)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (148948)..(148948)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (163385)..(163385)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (191989)..(191989)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (191995)..(191995)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (231980)..(231980)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (234187)..(234187)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (234220)..(234220)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (234814)..(234814)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (309398)..(309398)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (309418)..(309418)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (312837)..(312837)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (312993)..(312993)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (319226)..(319226)  
 OTHER INFORMATION: n equals a, t, c, or g  
 NAME/KEY: misc feature  
 LOCATION: (559167)..(559167)  
 OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc\_feature  
LOCATION: (559241)..(559241)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (600992)..(600992)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (622708)..(622708)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (657081)..(657081)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (657203)..(657203)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (674435)..(674435)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (682442)..(682442)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (713652)..(713652)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (741684)..(741684)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (779455)..(779455)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (779676)..(779676)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (855539)..(855539)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (871619)..(871619)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1084830)..(1084830)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1096846)..(1096846)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1119881)..(1119881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1130881)..(1130881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1310988)..(1310988)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1313224)..(1313224)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1349473)..(1349473)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1349491)..(1349491)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1470091)..(1470091)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1569020)..(1569020)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1602912)..(1602912)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature

LOCATION: (1603734)..(1603734)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1637998)..(1637998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1664854)..(1664854)  
OTHER INFORMATION: n equals a, t, c, or g  
US-08-916-421B-1

Query Match 2.0%; Score 21; DB 4; Length 1664976;  
Best Local Similarity 100.0%; Pred. No. 7.2;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 138 TGGTGGTGATGATGGTGG 158  
Db 568047 TGGTGGTGATGATGGTGG 568027

RESULT 14

US-09-769-482-9  
Sequence 9, Application US/09769482  
Patent No. 6566130  
GENERAL INFORMATION:  
APPLICANT: SRIVASTAVA, SHIV  
APPLICANT: MOUL, JUDD W.  
APPLICANT: XU, LINDA L.  
APPLICANT: SEGAWA, TAKEHIKO  
TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED  
TITLE OF INVENTION: POYNUCLEOTIDE ARRAY  
FILE REFERENCE: 04995.0057-00000  
CURRENT APPLICATION NUMBER: US/09/769,482  
CURRENT FILING DATE: 2001-01-26  
PRIOR APPLICATION NUMBER: 60/178,772  
PRIOR FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: 60/179,045  
PRIOR FILING DATE: 2000-01-31  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Primer  
US-09-769-482-9

Query Match 1.9%; Score 20; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 363 CCTTGGCCCGAGCGGAGCGC 382  
Db 1 CCTTGGCCCGAGCGGAGCGC 20

RESULT 15

US-08-753-247-22/c  
Sequence 22, Application US/08753247  
Patent No. 6210929  
GENERAL INFORMATION:  
APPLICANT: SCHLOKAT, Uwe  
APPLICANT: FISCHER, Bernhard  
APPLICANT: FALKNER, Falko-Guenther  
APPLICANT: DORNER, Friedrich  
APPLICANT: EIBL, Johann  
TITLE OF INVENTION: A FUSION PROTEIN COMPRISING A FURIN  
TITLE OF INVENTION: DERIVATIVE OR A DERIVATIVE OF A FURIN ANALOGUE AND A  
TITLE OF INVENTION: HETEROLOGOUS SEQUENCE  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500

CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20007-5109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/753,247  
 FILING DATE: 22-NOV-1996  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: AT 1928/95  
 FILING DATE: 24-NOV-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BENT, Stephen A.  
 REGISTRATION NUMBER: 29,768  
 REFERENCE/DOCKET NUMBER: 40433/149  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202)672-5300  
 TELEFAX: (202)672-5399  
 TELEX: 904136  
 INFORMATION FOR SEQ ID NO: 22:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 50 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 US-08-753-247-22

Query Match 1.9%; Score 20; DB 3; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 144 TGATGATGGTGGTGGTG 163  
 Db 42 TGATGATGGTGGTGGTG 23

Search completed: May 26, 2004, 20:27:26  
 Job time : 113 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 19:11:41 ; Search time 709 Seconds  
(without alignments)  
6801.066 Million cell updates/sec

Title: US-09-857-826B-44  
Perfect score: 1061  
Sequence: 1 tctctctgggttcgggtga.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: OLIGO NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 2953838 seqs, 2272363821 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 5907676

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	961	90.6	969	10	US-09-796-753-55
2	926	87.3	1141	15	US-10-205-823-412
3	926	87.3	1141	15	US-10-301-822-208
4	926	87.3	1850	15	US-10-241-220-44
5	926	87.3	4527	10	US-09-821-812-2
6	912	86.0	1140	15	US-10-390-045-1
7	912	86.0	1140	17	US-10-434-479-1
8	902	85.0	4839	15	US-10-241-220-119
9	902	85.0	4839	15	US-10-269-909-84
10	902	85.0	4839	15	US-10-269-909-85
11	891	84.0	1066	14	US-10-098-841-71
12	788	74.3	806	15	US-10-241-220-45
13	749	70.6	1321	9	US-09-934-249-1
14	708	66.7	759	15	US-10-390-045-2

15	708	66.7	759	17	US-10-434-479-2	Sequence 2, Appli
16	704	66.4	864	16	US-10-295-027-127	Sequence 127, App
17	701	66.1	861	9	US-09-934-249-3	Sequence 3, Appli
18	509	48.0	1583	15	US-10-000-256A-32	Sequence 32, Appli
19	313	29.5	693	9	US-09-934-249-14	Sequence 14, Appli
20	269	25.4	467	10	US-09-918-995-2074	Sequence 2074, Ap
21	155	14.6	168	13	US-10-085-783A-34699	Sequence 34699, A
22	155	14.6	168	16	US-10-242-535A-34699	Sequence 34699, A
23	129	12.2	225	13	US-10-085-783A-45532	Sequence 45532, A
24	129	12.2	225	16	US-10-242-535A-45532	Sequence 45532, A
25	102	9.6	368	9	US-09-783-590-3464	Sequence 3464, Ap
26	92	8.7	114	13	US-10-085-783A-31344	Sequence 31344, A
27	92	8.7	114	16	US-10-242-535A-31344	Sequence 31344, A
28	68	6.4	522	13	US-10-240-425-166	Sequence 166, App
29	60	5.7	60	10	US-09-908-975-13620	Sequence 13620, A
30	50	4.7	65	9	US-09-783-590-3488	Sequence 3488, Ap
31	44	4.1	878	9	US-09-934-249-12	Sequence 12, Appli
32	44	4.1	1713	10	US-09-796-753-57	Sequence 57, Appli
33	32	3.0	577	9	US-09-864-761-20542	Sequence 20542, A
34	32	3.0	1964	9	US-09-864-761-3776	Sequence 3776, Ap
35	29	2.7	401	9	US-09-864-761-3936	Sequence 3936, Ap
36	29	2.7	446	9	US-09-864-761-20699	Sequence 20699, A
37	29	2.7	475	9	US-09-934-249-15	Sequence 15, Appli
38	29	2.7	3444	15	US-10-293-582-16	Sequence 16, Appli
39	27	2.5	51	16	US-10-418-182-250	Sequence 250, App
40	26	2.5	1493	15	US-10-029-386-25133	Sequence 25133, A
41	26	2.5	7733	16	US-09-860-670-159	Sequence 159, App
42	26	2.5	7733	16	US-10-227-646-159	Sequence 159, App
43	26	2.5	8429	9	US-09-738-885-3	Sequence 3, Appli
44	26	2.5	439892	13	US-10-087-192-454	Sequence 454, App
45	25	2.4	331	9	US-09-864-761-17053	Sequence 17053, A

ALIGNMENTS

RESULT 1  
US-09-796-753-55  
; Sequence 55, Application US/09796753  
; Publication No. US20030027928A1  
; GENERAL INFORMATION:  
; APPLICANT: McCarthy, Sean A.  
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF  
; FILE REFERENCE: 7853-227-999  
; CURRENT APPLICATION NUMBER: US/09/796,753  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 09/183,175  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: 09/223,094  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/223,546  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/224,246  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/259,388  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR APPLICATION NUMBER: 60/122,458  
; PRIOR FILING DATE: 1999-03-01  
; PRIOR APPLICATION NUMBER: 09/312,359  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 09/336,536  
; PRIOR FILING DATE: 1999-06-18  
; PRIOR APPLICATION NUMBER: 09/342,687  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: 09/345,464  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: 09/365,164  
; PRIOR FILING DATE: 1999-07-30  
; PRIOR APPLICATION NUMBER: 09/399,723  
; PRIOR FILING DATE: 1999-09-20  
; PRIOR APPLICATION NUMBER: 09/409,634  
; PRIOR FILING DATE: 1999-09-30  
; PRIOR APPLICATION NUMBER: 09/471,179



US-10-205-823-412

```
Query Match      87.3%; Score 926; DB 15; Length 1141;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 976; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 81 GTCTCCTGCGAAACAGGCAATGGCGGAGCTGGAGTTGTTTCAGATCATCATCGTGG 140
Db 76 GTCTCCTGCGAAACAGGCAATGGCGGAGCTGGAGTTGTTTCAGATCATCATCGTGG 135

QY 141 TGGTGATGATGGTGATGGTGATGATCAGTGCTGCTGCTGAGCCACTACAAGCTGTCTG 200
Db 136 TGGTGATGATGGTGATGGTGATGATCAGTGCTGCTGCTGAGCCACTACAAGCTGTCTG 195

QY 201 CACGGTCCCTTCATCAGCCGGCACAGCCAGGGGCGGAGGAGAGAGATGCCCTGTCTCAG 260
Db 196 CACGGTCCCTTCATCAGCCGGCACAGCCAGGGGCGGAGGAGAGAGATGCCCTGTCTCAG 255

QY 261 AAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAAATCCAGAGCCGAGG 320
Db 256 AAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAAATCCAGAGCCGAGG 315

QY 321 TCTACGCCCGCCCTGCGGCCACCGACCGCTGGCCCTGGCCCGCCCTTCGCCCGGAGC 380
Db 316 TCTACGCCCGCCCTGCGGCCACCGACCGCTGGCCCGCCCTTCGCCCGGAGC 375

QY 381 GCTTCCACCGCTTCAGACCCACCTATCCGTACCTGAGCAGCAGATCGACCTGCCGCCCA 440
Db 376 GCTTCCACCGCTTCAGACCCACCTATCCGTACCTGAGCAGCAGATCGACCTGCCGCCCA 435

QY 441 CCATCTCGCTGTGACGGGAGGAGCCGCCACCTTACAGGGCCCTGCACCTCCAGC 500
Db 436 CCATCTCGCTGTGACGGGAGGAGCCGCCACCTTACAGGGCCCTGCACCTCCAGC 495

QY 501 TTCGGGACCCCGAGCAGCTGGAACTGAACCGGGAGTGGTGCGGCGACCCCAACA 560
Db 496 TTCGGGACCCCGAGCAGCTGGAACTGAACCGGGAGTGGTGCGGCGACCCCAACA 555

QY 561 GAACCATCTTCGACAGTGAACCTGATGGATAGTGCAGGCTGGGCGCCCTGCCCCCA 620
Db 556 GAACCATCTTCGACAGTGAACCTGATGGATAGTGCAGGCTGGGCGCCCTGCCCCCA 615

QY 621 GCAGTAACTCGGGCATCAGCGCCACGTGCTACGGCAGCGCGCGGCGCATGAGGGGCGCG 680
Db 616 GCAGTAACTCGGGCATCAGCGCCACGTGCTACGGCAGCGCGCGGCGCATGAGGGGCGCG 675

QY 681 CGCCACCTACAGCGAGGTTCATCGGCCACTACCGGGGTCTCTCTCCAGCAGCAGAGA 740
Db 676 CGCCACCTACAGCGAGGTTCATCGGCCACTACCGGGGTCTCTCTCCAGCAGCAGAGA 735

QY 741 GCAGTGGGCGCCCTCTCTGCTGGAGGGGACCGCGCTCCACCACACACATCGCGCC 800
Db 736 GCAGTGGGCGCCCTCTCTGCTGGAGGGGACCGCGCTCCACCACACACATCGCGCC 795

QY 801 TAGAGAGCGGAGCCATCTGGAGCAAGAGAGAGGATAAACAGAAAGGACACCTCTCTAGG 860
Db 796 TAGAGAGCGGAGCCATCTGGAGCAAGAGAGAGGATAAACAGAAAGGACACCTCTCTAGG 855

QY 861 GTCCCGAGGGGCGCGGCTGGGGTGGCTAGGTAGTGAAGGAGGAGAGTCCCGCTTCT 920
Db 856 GTCCCGAGGGGCGCGGCTGGGGTGGCTAGGTAGTGAAGGAGGAGAGTCCCGCTTCT 915

QY 921 TAGAAGAGGAGTGAGAGGAGGAGGGGGGCGGAGCAACGATCGTGTGGCCCTCCCTCC 980
Db 916 TAGAAGAGGAGTGAGAGGAGGAGGGGGGCGGAGCAACGATCGTGTGGCCCTCCCTCC 975

QY 981 CACCTCCCTGTGTATAAATATTATCATGTGATGTCTGGTCTGAATGCACAGCTAAGAGA 1040
Db 976 CACCTCCCTGTGTATAAATATTATCATGTGATGTCTGGTCTGAATGCACAGCTAAGAGA 1035

QY 1041 GCTTGCAAAAAA 1057
Db 1036 GCTTGCAAAAAA 1052
```

```
RESULT 3
US-10-301-822-208
; Sequence 208, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John B.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burt, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 208
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(854)
US-10-301-822-208
```

```
Query Match      87.3%; Score 926; DB 15; Length 1141;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 976; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 81 GTCTCCTGCGAAACAGGCAATGGCGGAGCTGGAGTTGTTTCAGATCATCATCGTGG 140
Db 76 GTCTCCTGCGAAACAGGCAATGGCGGAGCTGGAGTTGTTTCAGATCATCATCGTGG 135

QY 141 TGGTGATGATGGTGATGGTGATGATCAGTGCTGCTGCTGAGCCACTACAAGCTGTCTG 200
Db 136 TGGTGATGATGGTGATGGTGATGATCAGTGCTGCTGCTGAGCCACTACAAGCTGTCTG 195

QY 201 CACGGTCCCTTCATCAGCCGGCACAGCCAGGGGCGGAGGAGAGAGATGCCCTGTCTCAG 260
Db 196 CACGGTCCCTTCATCAGCCGGCACAGCCAGGGGCGGAGGAGAGAGATGCCCTGTCTCAG 255

QY 261 AAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAAATCCAGAGCCGAGG 320
Db 256 AAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAAATCCAGAGCCGAGG 315

QY 321 TCTACGCCCGCCCTGCGGCCACCGACCGCTGGCCCTGGCCCGCCCTTCGCCCGGAGC 380
Db 316 TCTACGCCCGCCCTGCGGCCACCGACCGCTGGCCCGCCCTTCGCCCGGAGC 375

QY 381 GCTTCCACCGCTTCAGACCCACCTATCCGTACCTGAGCAGCAGATCGACCTGCCGCCCA 440
Db 376 GCTTCCACCGCTTCAGACCCACCTATCCGTACCTGAGCAGCAGATCGACCTGCCGCCCA 435

QY 441 CCATCTCGCTGTGACGGGAGGAGCCGCCACCTTACAGGGCCCTGCACCTCCAGC 500
Db 436 CCATCTCGCTGTGACGGGAGGAGCCGCCACCTTACAGGGCCCTGCACCTCCAGC 495

QY 501 TTCGGGACCCCGAGCAGCTGGAACTGAACCGGGAGTGGTGCGGCGACCCCAACA 560
Db 496 TTCGGGACCCCGAGCAGCTGGAACTGAACCGGGAGTGGTGCGGCGACCCCAACA 555
```





```

; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4527
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(851)
US-09-821-812-2

```

	Query Match	87.3%;	Score 926;	DB 10;	Length 4527;
	Best Local Similarity	99.9%;	Pred. No. 0;		
	Matches 976;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	81	GTCTCTCGAACCAGGCAATGGCGGAGCTGGAGTTTGTTTTCAGATCATCATCATCTGCTGG	140		
DB	76	GTCTCTCGAACCAGGCAATGGCGGAGCTGGAGTTTGTTTTCAGATCATCATCATCTGCTGG	135		
QY	141	TGGTGATGATGGTGTGATGGTGGTGGTGTGATCACGTCCTGCTGAGCCACTACAAGCTGTCTG	200		
DB	136	TGGTGATGATGGTGTGATGGTGGTGGTGTGATCACGTCCTGCTGAGCCACTACAAGCTGTCTG	195		
QY	201	CACGGTCTTCATCAGCCCGGCACAGCCAGGGCGGAGAGAGAAGATGCCCTGTCTCTCAG	260		
DB	196	CACGGTCTTCATCAGCCCGGCACAGCCAGGGCGGAGAGAGAAGATGCCCTGTCTCTCAG	255		
QY	261	AAGGATGCCTGTGGCCTTCGGAGAGCACAGTGTACGGCAACGGAAATCCAGAGCCGCAGG	320		
DB	256	AAGGATGCCTGTGGCCTTCGGAGAGCACAGTGTACGGCAACGGAAATCCAGAGCCGCAGG	315		
QY	321	TCTACGCCCCCGCTCGGCCACCGACCGCTGGCCGTGGCCGCCCTTCGCCCCAGCGGAGC	380		
DB	316	TCTACGCCCCCGCTCGGCCACCGACCGCTGGCCGTGGCCGCCCTTCGCCCCAGCGGAGC	375		
QY	381	GCTTCCACCGCTTCAGCCCCACCTATCCGTACCTGCAGACAGAGATCGACCTGCCGCCCA	440		
DB	376	GCTTCCACCGCTTCAGCCCCACCTATCCGTACCTGCAGACAGAGATCGACCTGCCGCCCA	435		
QY	441	CCATCTCGCTGTACACGGGGAGGAGCCCCCAACCTACAGGGGCCCTGCACCCCTCCAGC	500		
DB	436	CCATCTCGCTGTACACGGGGAGGAGCCCCCAACCTACAGGGGCCCTGCACCCCTCCAGC	495		
QY	501	TTCGGACCCCGAGCAGAGCTGGAACTGAACCGGGAGTGGTGCAGCGCACCCGCCAAACA	560		
DB	496	TTCGGACCCCGAGCAGAGCTGGAACTGAACCGGGAGTGGTGCAGCGCACCCGCCAAACA	555		
QY	561	GAACCATCTTCGACAGTACCTGATGGATAGTGCAGGCTGGCGGCCCTTGCCTCCGCCCA	620		
DB	556	GAACCATCTTCGACAGTACCTGATGGATAGTGCAGGCTGGCGGCCCTTGCCTCCGCCCA	615		
QY	621	GCAGTAACTCGGGCATCAGCGCCACGTGCTACGGGCAGCGCGGGCGGCATGGAGGGGCGC	680		
DB	616	GCAGTAACTCGGGCATCAGCGCCACGTGCTACGGGCAGCGCGGGCGGCATGGAGGGGCGC	675		
QY	681	CGCCCCACTACAGCGAGGTCTCGGCCACTACCCGGGGTCTCTCTCCAGCACCGACAGA	740		
DB	676	CGCCCCACTACAGCGAGGTCTCGGCCACTACCCGGGGTCTCTCTCCAGCACCGACAGA	735		
QY	741	GCAGTGGGCCGCCCTCTCTGCTGGAGGGGACCCGGCTCCACACACACATCGCGCCCC	800		
DB	736	GCAGTGGGCCGCCCTCTCTGCTGGAGGGGACCCGGCTCCACACACACATCGCGCCCC	795		
QY	801	TAGAGAGCGCAGCCATCTGGAGCAAAGAGAAGGATATAACAGAAAGGACACCTCTCTAGG	860		
DB	796	TAGAGAGCGCAGCCATCTGGAGCAAAGAGAAGGATATAACAGAAAGGACACCTCTCTAGG	855		
QY	861	GTCCCCCAGGGGGCCCGGCTGGGGCTGGTAGGTGAAAAGGCAGAACACTCCGGCGTTCT	920		
DB	856	GTCCCCCAGGGGGCCCGGCTGGGGCTGGTAGGTGAAAAGGCAGAACACTCCGGCGTTCT	915		
QY	921	TAGAAGAGGATGAGAGGAAGCGGGGGGGCGAGCAACGCATCGTGTGGGCCCTCCCTCC	980		

Db	916		TAGAGAGGAGTGAGAGGAAGGGGGGGCCAGCAACGCATCGTGTGGCCCTCCCTCC	975
QY	981		CACCTCCCTGCTATAAATATTTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGA	1040
Db	976		CACCTCCCTGCTATAAATATTTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGA	1035
QY	1041		GCTTGCAAAAAAAAAA	1057
Db	1036		GCTTGCAAAAAAAAAA	1052

```

RESULT 6
US-10-390-045-1
; Sequence 1, Application US/10390045
; Publication No. US20030170713A1
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/10/390,045
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US/09/769,482
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)..(850)
US-10-390-045-1

```

	Query Match	86.0%;	Score 912;	DB 15;	Length 1140;
	Best Local Similarity	99.9%;	Pred. No. 0;		
	Matches 962;	Conservative	0;	Mismatches	1; Indels 0; Gaps 0;
QY	95	CAGGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGCATGATGGTG	154		
DB	89	CAGGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGCATGATGGTG	148		
QY	155	ATGGTGGTGTGATCAGCTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCCCTTCATC	214		
DB	149	ATGGTGGTGTGATCAGCTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCCCTTCATC	208		
QY	215	AGCCGGCACAGCCAGGGGCGGAGGAGAGATGCCCTGTCTCAGAAGGATGCCCTGTGG	274		
DB	209	AGCCGGCACAGCCAGGGGCGGAGGAGAGATGCCCTGTCTCAGAAGGATGCCCTGTGG	268		
QY	275	CCCTCGGAGAGCACAGTGT CAGGCAACGGAAATCCAGAGCCGCAGGTCTACGCCCCGCT	334		
DB	269	CCCTCGGAGAGCACAGTGT CAGGCAACGGAAATCCAGAGCCGCAGGTCTACGCCCCGCT	328		
QY	335	CGGCCCCACCGACCGCCTGGCCGCTGCCGCCCTTGCCCGAGCGGGAGCGCTTCCACCGCTTC	394		
DB	329	CGGCCCCACCGACCGCCTGGCCGCTGCCGCCCTTGCCCGAGCGGGAGCGCTTCCACCGCTTC	388		
QY	395	CAGCCCCACCTATCCGTACTCTGCAGCACGAGATCGACCTGCCGCCCCACCATCTCGCTGTCA	454		
DB	389	CAGCCCCACCTATCCGTACTCTGCAGCACGAGATCGACCTGCCGCCCCACCATCTCGCTGTCA	448		
QY	455	GACGGGGAGGAGCCCCACCCCTACGAGGGGCCCTGCACCTCCAGCTTCGGGACCCCGAG	514		



Db 449 GACGGGAGAGCCCCACCCCTACCAAGGSCCCCTGACCCCTCCAGCTTCGGGAGCCCGAG 508  
QY 515 CAGCAGCTGGAACCTGAACCGGAGTCGGTGGCGGCACCCGCCAAACAGAACCATCTTCGAC 574  
Db 509 CAGCAGCTGGAACCTGAACCGGAGTCGGTGGCGGCACCCGCCAAACAGAACCATCTTCGAC 568  
QY 575 AGTGACCTGATGATAGTGTCCAGGCTGGGCGGCCCTGCCCCCAGCAGTAACTCGGGC 634  
Db 569 AGTGACCTGATGATAGTGTCCAGGCTGGGCGGCCCTGCCCCCAGCAGTAACTCGGGC 628  
QY 635 ATCAGGCGCACGTCGTACGGCAGCGGCGCGCATGGAGGGCGCGCCACCTACAGC 694  
Db 629 ATCAGGCGCACGTCGTACGGCAGCGGCGCGCATGGAGGGCGCGCCACCTACAGC 688  
QY 695 GAGGTCTATCGCCACTACCCGGGCTCCCTCCAGCACACACATCGGCGCCCTAGAGAGCGCAGCC 754  
Db 689 GAGGTCTATCGCCACTACCCGGGCTCCCTCCAGCACACACATCGGCGCCCTAGAGAGCGCAGCC 748  
QY 755 TCCTTGCTGAGGGGACCCGGCTCCACCAACACACATCGGCGCCCTAGAGAGCGCAGCC 814  
Db 749 TCCTTGCTGAGGGGACCCGGCTCCACCAACACACATCGGCGCCCTAGAGAGCGCAGCC 808  
QY 815 ATCTGAGCAAGAGAGGATAAACAGAAAGGACACCTCTCTAGGTCCTCCAGGGGGC 874  
Db 809 ATCTGAGCAAGAGAGGATAAACAGAAAGGACACCTCTCTAGGTCCTCCAGGGGGC 868  
QY 875 CGGGCTGGGCTCGTAGGTGAAAGGAGCAACACTCCGGCGCTCTTAGAAGAGGAGTGA 934  
Db 869 CGGGCTGGGCTCGTAGGTGAAAGGAGCAACACTCCGGCGCTCTTAGAAGAGGAGTGA 928  
QY 935 GAGGAAGCGGGGGCGCAGCAACGCTATGTTGGCCCTCCCTCCAGCTCCCTGTGTA 994  
Db 929 GAGGAAGCGGGGGCGCAGCAACGCTATGTTGGCCCTCCCTCCAGCTCCCTGTGTA 988  
QY 995 TAAATATTACATGTGATGTCTGGTCTGAATGCAAGCTAAGAGAGCTTGCAAAAAA 1054  
Db 989 TAAATATTACATGTGATGTCTGGTCTGAATGCAAGCTAAGAGAGCTTGCAAAAAA 1048  
QY 1055 AAA 1057  
Db 1049 AAA 1051

RESULT 7  
US-10-434-479-1  
; Sequence 1, Application US/10434479  
; Publication No. US2004092469A1  
; GENERAL INFORMATION:  
; APPLICANT: SRIVASTAVA, SHIV  
; APPLICANT: MOUL, JUDD W.  
; APPLICANT: XU, LINDA L.  
; TITLE OF INVENTION: ANDROGEN-REGULATED PMEPA1 GENE AND POLYPEPTIDES  
; FILE REFERENCE: 04995.0057-0200  
; CURRENT APPLICATION NUMBER: US/10/434,479  
; CURRENT FILING DATE: 2003-05-09  
; PRIOR APPLICATION NUMBER: 10/390,045  
; PRIOR FILING DATE: 2003-03-18  
; PRIOR APPLICATION NUMBER: 09/769,482  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,772  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/179,045  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1140  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (95)..(850)  
US-10-434-479-1

Query Match 86.0%; Score 912; DB 17; Length 1140;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 962; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 95 CAGGCAATCGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGAATGGTG 154  
Db 89 CAGGCAATCGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGAATGGTG 148  
QY 155 ATGGTGGTGGTGAATCAAGCTGCTGAGCACTAAGCTGTCTGACGCTCTTCATC 214  
Db 149 ATGGTGGTGGTGAATCAAGCTGCTGAGCACTAAGCTGTCTGACGCTCTTCATC 208  
QY 215 AGCCGGCAAGCCAGGGGGGAGGAGAGATGCCCTGTCTCAGAGGATGCTGTGG 274  
Db 209 AGCCGGCAAGCCAGGGGGGAGGAGAGATGCCCTGTCTCAGAGGATGCTGTGG 268  
QY 275 CCCTCGGAGAGCAGTGTAGGCAACGGAAATCCAGAGCCGAGGTCTACGCCCCGCT 334  
Db 269 CCCTCGGAGAGCAGTGTAGGCAACGGAAATCCAGAGCCGAGGTCTACGCCCCGCT 328  
QY 335 CGGCCCCACGACCGCTGGCCGTGCCCTTCGCGCCAGCGGAGCGCTTCACCGCTTC 394  
Db 329 CGGCCCCACGACCGCTGGCCGTGCCCTTCGCGCCAGCGGAGCGCTTCACCGCTTC 388  
QY 395 CAGCCCCACCTATCGTACCTGACGACGAGATCGACCTGCCGCCACCATCTCGCTGCA 454  
Db 389 CAGCCCCACCTATCGTACCTGACGACGAGATCGACCTGCCGCCACCATCTCGCTGCA 448  
QY 455 GACGGGGAGAGCCCCCACCCTACAGGGSCCTTGACCCCTCCAGCTTCGGGAGCCCGAG 514  
Db 449 GACGGGGAGAGCCCCCACCCTACAGGGSCCTTGACCCCTCCAGCTTCGGGAGCCCGAG 508  
QY 515 CAGCAGCTGGAACCTGAACCGGGAGTCGGTGGCGGCAACCCCAACAGAACCATCTTCGAC 574  
Db 509 CAGCAGCTGGAACCTGAACCGGGAGTCGGTGGCGGCAACCCCAACAGAACCATCTTCGAC 568  
QY 575 AGTGACCTGATGATAGTGTCCAGGCTGGGCGGCCCTGCCCCCAGCAGTAACTCGGGC 634  
Db 569 AGTGACCTGATGATAGTGTCCAGGCTGGGCGGCCCTGCCCCCAGCAGTAACTCGGGC 628  
QY 635 ATCAGCGCCAGCTGTCTACGGCAGCGGCGGCGCATGGAGGGCGCCGCCCTACCTACAGC 694  
Db 629 ATCAGCGCCAGCTGTCTACGGCAGCGGCGGCGCATGGAGGGCGCCGCCCTACCTACAGC 688  
QY 695 GAGGTCTATCGGCGCACTACCCGGGGTCTCTTCAGCAGCAGCAGAGCAGTGGGCGGCC 754  
Db 689 GAGGTCTATCGGCGCACTACCCGGGGTCTCTTCAGCAGCAGCAGCAGAGCAGTGGGCGGCC 748  
QY 755 TCCTTGCTGAGGGGACCCGGCTCCACCACACACACATCGGCGCCCTAGAGAGCGCAGCC 814  
Db 749 TCCTTGCTGAGGGGACCCGGCTCCACCACACACACATCGGCGCCCTAGAGAGCGCAGCC 808  
QY 815 ATCTGAGCAAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGC 874  
Db 809 ATCTGAGCAAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGC 868  
QY 875 CGGGCTGGGCTCGTAGGTGAAAGGAGGAGCACTCGTGGCGCCCTCCCTGTGTA 934  
Db 869 CGGGCTGGGCTCGTAGGTGAAAGGAGGAGCACTCGTGGCGCCCTCCCTGTGTA 928  
QY 935 GAGGAAGCGGGGGCGCAGCAACGCTATGTTGGCCCTCCCTCCAGCTCCCTGTGTA 994  
Db 929 GAGGAAGCGGGGGCGCAGCAACGCTATGTTGGCCCTCCCTCCAGCTCCCTGTGTA 988  
QY 995 TAAATATTACATGTGATGTCTGGTCTGAATGCAAGCTAAGAGAGCTTGCAAAAAA 1054  
Db 989 TAAATATTACATGTGATGTCTGGTCTGAATGCAAGCTAAGAGAGCTTGCAAAAAA 1048  
QY 1055 AAA 1057  
Db 1049 AAA 1051



```

RESULT 8
US-10-241-220-119
; Sequence 119, Application US/10241220
; Publication No. US20030148408A1
; GENERAL INFORMATION:
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Phillips, Heidi
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TREATMENT OF TUMOR
; FILE REFERENCE: P5010R1-US
; CURRENT APPLICATION NUMBER: US/10/241,220
; CURRENT FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 119
; LENGTH: 4839
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-241-220-119

Query Match      85.0%; Score 902; DB 15; Length 4839;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 952; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 105 CGGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGGTGGTGGTGGTGG 164
Db 430 CGGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGGTGGTGGTGG 489

QY 165 TGATCAGTGCCTGCTGAGCCACTACAGCTGTGTGACGGTCTTTCATCAGCCGGCACA 224
Db 490 TGATCAGTGCCTGCTGAGCCACTACAGCTGTGTGACGGTCTTTCATCAGCCGGCACA 549

QY 225 GCCAGGGGGGAGGAGAGAGATGCCCTGTCTCAGAGGATGCTTGTGGCCCTCGGAGA 284
Db 550 GCCAGGGGGGAGGAGAGAGATGCCCTGTCTCAGAGGATGCTTGTGGCCCTCGGAGA 609

QY 285 GCACAGTGTGAGCAACGGAAATCCAGAGCCGAGGTCTACGCCCGCCCTCGGCCACCG 344
Db 610 GCACAGTGTGAGCAACGGAAATCCAGAGCCGAGGTCTACGCCCGCCCTCGGCCACCG 669

QY 345 ACCGCTGCGCGTGGCCGCTTGGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACCT 404
Db 670 ACCGCTGCGCGTGGCCGCTTGGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACCT 729

QY 405 ATCCGCTACCTGACGACGAGATCGACCTGCGCCCAACCATCTCGCTGTGACAGCGGGAGG 464
Db 730 ATCCGCTACCTGACGACGAGATCGACCTGCGCCCAACCATCTCGCTGTGACAGCGGGAGG 789

QY 465 AGCCCCACCTTACAGGGGGCCCTGCACTTCCAGCTTCCAGCCCGGAGCGAGCTGG 524
Db 790 AGCCCCACCTTACAGGGGGCCCTGCACTTCCAGCTTCCAGCCCGGAGCGAGCTGG 849

QY 525 AACTGAACCGGAGTGGTGGCGGACCCCAACAGAAACCATCTTCAGAGTGACCTGA 584
Db 850 AACTGAACCGGAGTGGTGGCGGACCCCAACAGAAACCATCTTCAGAGTGACCTGA 909

QY 585 TGGATAGTGCAGGCTGGGGGGCCCTGCGCCCGCCAGCAGTAACCTCGGCGATCAGGCCA 644
Db 910 TGGATAGTGCAGGCTGGGGGGCCCTGCGCCCGCCAGCAGTAACCTCGGCGATCAGGCCA 969

QY 645 CGTGCTACGCGAGCGGGGGCGATGGAGGGGCGCGCCCAACCTACAGCGAGTCACTCG 704
Db 970 CGTGCTACGCGAGCGGGGGCGATGGAGGGGCGCGCGCCCAACCTACAGCGAGTCACTCG 1029

QY 705 GCCACTACCGGGTCTCTTCCAGCAACAGCAGAGCAGTGGGGCGCCCTCTCTTGTGG 764
Db 1030 GCCACTACCGGGTCTCTTCCAGCAACAGCAGAGCAGTGGGGCGCCCTCTCTTGTGG 1089

```

```

QY 765 AGGGACCGGCTCCACCACACACATCGGCGCCCTAGAGAGCGCAGCCATCTGGAGCA 824
Db 1090 AGGGACCGGCTCCACCACACACATCGGCGCCCTAGAGAGCGCAGCCATCTGGAGCA 1149

QY 825 AAGAGAAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGGGCGGCTGGGG 884
Db 1150 AAGAGAAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGGGCGGCTGGGG 1209

QY 885 CTGCGTAGGTGAAAAGGCAAGACACTCCGCGCTTCTTAGAAGAGAGTGAGAGGAGCGG 944
Db 1210 CTGCGTAGGTGAAAAGGCAAGACACTCCGCGCTTCTTAGAAGAGAGTGAGAGGAGCGG 1269

QY 945 GGGGGCGCAGCAACGCAATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATTTA 1004
Db 1270 GGGGGCGCAGCAACGCAATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATTTA 1329

QY 1005 CATGTGATGTCTGCTGAATGCAAGCTAAGAGAGCTTTGCAAAAAA 1057
Db 1330 CATGTGATGTCTGCTGAATGCAAGCTAAGAGAGCTTTGCAAAAAA 1382

RESULT 9
US-10-269-909-84
; Sequence 84, Application US/10269909
; Publication No. US20030180747A1
; GENERAL INFORMATION:
; APPLICANT: HRUBAN, RALPH H.
; APPLICANT: ARGANI, PEDRAM
; APPLICANT: IACOBUZIO-DONAHUE, CHRISTINE
; APPLICANT: MAITRA, ANIRBAN
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
; FILE REFERENCE: 58303(71699)
; CURRENT APPLICATION NUMBER: US/10/269,909
; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/332,754
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 84
; LENGTH: 4839
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-269-909-84

Query Match      85.0%; Score 902; DB 15; Length 4839;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 952; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 105 CGGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGGTGGTGGTGG 164
Db 430 CGGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGGTGGTGGTGG 489

QY 165 TGATCAGTGCCTGCTGAGCCACTACAGCTGTCTGACAGGTCCTTTCATCAGCCGGCACA 224
Db 490 TGATCAGTGCCTGCTGAGCCACTACAGCTGTCTGACAGGTCCTTTCATCAGCCGGCACA 549

QY 225 GCCAGGGGGGAGGAGAGAGATGCCCTGTCTCAGAGGATGCTTGTGGCCCTCGGAGA 284
Db 550 GCCAGGGGGGAGGAGAGAGATGCCCTGTCTCAGAGGATGCTTGTGGCCCTCGGAGA 609

QY 285 GCACAGTGTGAGCAACGGAAATCCAGAGCCGAGGTCTACGCCCGCCCTCGGCCACCG 344
Db 610 GCACAGTGTGAGCAACGGAAATCCAGAGCCGAGGTCTACGCCCGCCCTCGGCCACCG 669

QY 345 ACCGCTGCGCGTGGCCGCTTGGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACCT 404
Db 670 ACCGCTGCGCGTGGCCGCTTGGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACCT 729

QY 405 ATCCGCTACCTGACGACGAGATCGACCTGCGCCCAACCATCTCGCTGTGACAGCGGGAGG 464
Db 730 ATCCGCTACCTGACGACGAGATCGACCTGCGCCCAACCATCTCGCTGTGACAGCGGGAGG 789

QY 465 AGCCCCACCTTACAGGGGGCCCTGCACTTCCAGCTTCCAGCCCGGAGCGAGCTGG 524
Db 790 AGCCCCACCTTACAGGGGGCCCTGCACTTCCAGCTTCCAGCCCGGAGCGAGCTGG 849

QY 525 AACTGAACCGGAGTGGTGGCGGACCCCAACAGAAACCATCTTCAGAGTGACCTGA 584
Db 850 AACTGAACCGGAGTGGTGGCGGACCCCAACAGAAACCATCTTCAGAGTGACCTGA 909

QY 585 TGGATAGTGCAGGCTGGGGGGCCCTGCGCCCGCCAGCAGTAACCTCGGCGATCAGGCCA 644
Db 910 TGGATAGTGCAGGCTGGGGGGCCCTGCGCCCGCCAGCAGTAACCTCGGCGATCAGGCCA 969

QY 645 CGTGCTACGCGAGCGGGGGCGATGGAGGGGCGCGCCCAACCTACAGCGAGTCACTCG 704
Db 970 CGTGCTACGCGAGCGGGGGCGATGGAGGGGCGCGCGCCCAACCTACAGCGAGTCACTCG 1029

QY 705 GCCACTACCGGGTCTCTTCCAGCAACAGCAGAGCAGTGGGGCGCCCTCTCTTGTGG 764
Db 1030 GCCACTACCGGGTCTCTTCCAGCAACAGCAGAGCAGTGGGGCGCCCTCTCTTGTGG 1089

```









Db	193	AAGGATGCTGTGGCCCTCGGAGAGCAGAGTGT	CAGGCAACGGAATCCAGAGCCG	CAGG	255
Qy	321	TCTACGCCCCGCTCGGCCCAACCGACCGCTGGCCGTGCCGCCCTTCGCCCAAGCGGGAGC			380
Db	253	TCTACGCCCCGCTCGGCCCAACCGACCGCTGGCCGTGCCGCCCTTCGCCCAAGCGGGAGC			312
Qy	381	GCTTCCACCGCTTCCAGCCCACTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCCA			440
Db	313	GCTTCCACCGCTTCCAGCCCACTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCCA			372
Qy	441	CCATCTCGCTGTCCAGCGGGAGGAGCCCCCAACCTTACGAGGGCCCCCTGCACCTTCAGC			500
Db	373	CCATCTCGCTGTCCAGCGGGAGGAGCCCCCAACCTTACGAGGGCCCCCTGCACCTTCAGC			432
Qy	501	TTCGGGACCCCGAGCAGCAGCTGGAACTGAACCGGGAGTCGGTGGCGGCACCCCCAAACA			560
Db	433	TTCGGGACCCCGAGCAGCAGCTGGAACTGAACCGGGAGTCGGTGGCGGCACCCCCAAACA			492
Qy	561	GAACCATCTTCGACAGTGACCTGATGGATAGTGCAGGCTGGCGGGCCCCCTGCCCCCCCCA			620
Db	493	GAACCATCTTCGACAGTGACCTGATGGATAGTGCAGGCTGGCGGGCCCCCTGCCCCCCCCA			552
Qy	621	GCAGTAACTCGGGCATCAGCGCCACTGCTCTACGGCAGCGGGCGCATGAGAGGGCCGC			680
Db	553	GCAGTAACTCGGGCATCAGCGCCACTGCTCTACGGCAGCGGGCGCATGAGAGGGCCGC			612
Qy	681	CGCCACCTACAGCGAGGTCAFCGGCCACTACCCGGGGTCTCTTCCAGCACCCAGCAGA			740
Db	613	CGCCACCTACAGCGAGGTCAFCGGCCACTACCCGGGGTCTCTTCCAGCACCCAGCAGA			672
Qy	741	GCAGTGGGCGGCCCTCTTGTCTGGAGGGGACCCGGCTCCACACACACATCGCGCCCC			800
Db	673	GCAGTGGGCGGCCCTCTTGTCTGGAGGGGACCCGGCTCCACACACACATCGCGCCCC			732
Qy	801	TAGAGAGCGCAGCCATCTGGAGCAAGAGAAGGATAAACAGAAAGGACACCTCTCTAGG			860
Db	733	TAGAGAGCGCAGCCATCTGGAGCAAGAGAAGGATAAACAGAAAGGACACCTCTCTAGG			792
Qy	861	GTCCCCCAG			868
Db	793	GTCCCCCAG			800

**RESULT 13**

```

US-09-934-249-1
; Sequence 1, Application US/09934249
; Patent No. US20020115081A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Richard T.
; APPLICANT: Landschulz, Katherine T.
; APPLICANT: Turi, Thomas G.
; APPLICANT: Thompson, John P.
; APPLICANT: Kennedy, Scott P.
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; FILE REFERENCE: P0738/7001/ERP/KA
; CURRENT APPLICATION NUMBER: US/09/934,249
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,159
; PRIOR FILING DATE: 2000-08-22
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1321
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (413) ... (1273)
US-09-934-249-1

```

Query Match	70.6%;	Score 749;	DB 9;	Length 1321;
Best Local Similarity	99.9%;	Pred. No. 0;		

	Matches	799; Conservative	0; Mismatches	1; Indels	0; Gaps	0;
QY	105	CGGAGCTGGAGTTTGTTCAGATCATCATCTCGTGTTGTTGAATGTTGTTGGTGG	164			
DB	522	CGGAGCTGGAGTTTGTTCAGATCATCATCTCGTGTTGTTGAATGTTGTTGGTGG	581			
QY	165	TGATCACGTGCTCTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGGCACA	224			
DB	582	TGATCAOCTGCTGAGCCAATAAAGCTGTCTGCACGGTCTTCATCAGCCGGCACA	641			
QY	225	GCCAGGGGCGGAGGAGAGAAGATGCCCTGTCTCTCAGAAGGATGCCCTGCGCCCTCGGAGA	284			
DB	642	GCCAGGGGCGGAGGAGAGAAGATGCCCTGTCTCTCAGAAGGATGCCCTGCGCCCTCGGAGA	701			
QY	285	GCACAGTGTGAGGCAAACGGAATCCAGAGCCGAGGTCTACGCCCGCTCGGCCCCACCG	344			
DB	702	GCACAGTGTGAGGCAAACGGAATCCAGAGCCGAGGTCTACGCCCGCTCGGCCCCACCG	761			
QY	345	ACCGCTTGGCGTCCGCCCTTCGCCAGCGGAGCGCTTCACCGCTTCAGCCCCACCT	404			
DB	762	ACCGCTTGGCGTCCGCCCTTCGCCAGCGGAGCGCTTCACCGCTTCAGCCCCACCT	821			
QY	405	ATCCGTACCTGCAGCACGAGATCGACCTGCGGCCACCATTCTCGCTGTCTCAGACGGGAGG	464			
DB	822	ATCCGTACCTGCAGCACGAGATCGACCTGCGGCCACCATTCTCGCTGTCTCAGACGGGAGG	881			
QY	465	AGCCCCACCTACCCAGGGCCCCTGCAACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGG	524			
DB	882	AGCCCCACCTACCCAGGGCCCCTGCAACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGG	941			
QY	525	AACCTGAACCGGAGTTCGTGCGGCAACCCCAAACAGAACCATCTTCGACAGTGACCTGA	584			
DB	942	AACCTGAACCGGAGTTCGTGCGGCAACCCCAAACAGAACCATCTTCGACAGTGACCTGA	1001			
QY	585	TGGATAGTGCAGGCTGGGGCGGCATGAGGGGGCGCCGCCACCTACAGCGAGTCAACGCGCA	644			
DB	1002	TGGATAGTGCAGGCTGGGGCGGCATGAGGGGGCGCCGCCACCTACAGCGAGTCAACGCGCA	1061			
QY	645	CCTGTCTACGGCAGCGGGCGGCATGAGGGGGCGCCGCCACCTACAGCGAGTCAACGCGCA	704			
DB	1062	CCTGTCTACGGCAGCGGGCGGCATGAGGGGGCGCCGCCACCTACAGCGAGTCAACGCGCA	1121			
QY	705	GCCACTACCGGGGTCTCTTCCAGCACACACATCGCGCCCTTAGAGAGCGCAGCATCTGGAGCA	764			
DB	1122	GCCACTACCGGGGTCTCTTCCAGCACACACATCGCGCCCTTAGAGAGCGCAGCATCTGGAGCA	1181			
QY	765	AGGGGACCCGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCATCTGGAGCA	824			
DB	1182	AGGGGACCCGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCATCTGGAGCA	1241			
QY	825	AAGAGAAAGGATAAACAAGAACACCCCTCTCTAGGGTCCCGAGGGGGCGGGGCTGGGG	884			
DB	1242	AAGAGAAAGGATAAACAAGAACACCCCTCTCTAGGGTCCCGAGGGGGCGGGGCTGGGG	1301			
QY	885	CTGCGTAGGTGAAAAGGCAG	904			
DB	1302	CTGCGTAGGTGAAAAGGCAG	1321			

## RESULT 14

US-10-390-045-2  
; Sequence 2, Application US/10390045  
; Publication No. US20030170713A1  
; GENERAL INFORMATION:  
; APPLICANT: SRIVASTAVA, SHIV  
; APPLICANT: MOUL, JUDD W.  
; APPLICANT: XU, LINDA L.  
; APPLICANT: SEGAWA, TAKEHIKO  
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED  
; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY  
; FILE REFERENCE: 04995.0057-00000  
; CURRENT APPLICATION NUMBER: US/10/390,045  
; CURRENT FILING DATE: 2003-03-18



Qy 761 CTGGAGGGGACCCGGCTCCACCACACACATCGCGCCCTAGAGAGCGGAGCCATCTGG 820  
Db |||||

Qy 661 CTGGAGGGGACCCGGCTCCACCACACACATCGCGCCCTAGAGAGCGGAGCCATCTGG 720  
Db |||||

Qy 821 AGCAAAGAGAGGATATAACAGAAAGGACACCCCTCTCTAG 859  
Db |||||

Qy 721 AGCAAAGAGAGGATATAACAGAAAGGACACCCCTCTCTAG 759  
Db |||||

Search completed: May 26, 2004, 22:14:31  
Job time : 712 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 16:46:18 ; Search time 4416 Seconds  
(without alignments)  
7174.768 Million cell updates/sec

Title: US-09-857-826B-44  
Perfect score: 1061  
Sequence: 1 tcctcttggttcgggtga.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: OLIGO\_NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	722	68.0	729	13 BQ575741	BQ575741 UI-H-EZ1-
2	665	62.7	967	13 BQ641849	BQ641849 AGENCOURT
3	652	61.5	655	13 BQ691705	BQ691705 AGENCOURT
C 4	652	61.5	951	9 AL558881	AL558881 AL558881

5	609	57.4	609	13	BQ636742	BQ636742 hdl3h06.y
6	601	56.6	890	13	BQ690750	BQ690750 AGENCOURT
C 7	599	56.5	602	14	CA431191	CA431191 UI-H-FG1-
C 8	585	55.1	588	13	BU624784	BU624784 UI-H-FG1-
C 9	570	53.7	629	13	BU730650	BU730650 UI-B-CI1-
C 10	559	52.7	559	10	BE855409	BE855409 7g13f05.x
C 11	556	52.4	626	12	BM974296	BM974296 UI-CF-EC1
C 12	536	50.5	570	13	BQ575582	BQ575582 UI-H-EZ1-
13	536	50.5	945	13	BU539219	BU539219 AGENCOURT
14	530	50.0	563	14	CB049800	CB049800 NISC_gj13
15	528	49.8	844	13	BQ686793	BQ686793 AGENCOURT
16	528	49.8	952	13	BU157959	BU157959 AGENCOURT
17	525	49.5	572	13	BM641317	BM641317 DKFZp686K
C 18	523	49.3	730	12	BM677602	BM677602 UI-B-E01-
19	508	47.9	1046	12	BM922276	BM922276 AGENCOURT
20	505	47.6	646	13	BU859841	BU859841 AGENCOURT
21	499	47.0	551	12	BM141979	BM141979 if25a11.y
22	492	46.4	633	12	BM714472	BM714472 UI-B-EJ0-
C 23	491	46.3	782	12	BQ015170	BQ015170 UI-H-ED1-
C 24	484	45.6	547	12	BM676516	BM676516 UI-B-EJ0-
C 25	484	45.6	552	12	BM713900	BM713900 UI-B-EJ0-
C 26	480	45.2	728	13	BU683523	BU683523 UI-CF-EC1
27	476	44.9	1007	9	AL558882	AL558882 AL558882
28	469	44.2	613	12	BG680325	BG680325 602629217
29	468	44.1	964	13	BU859860	BU859860 AGENCOURT
30	461	43.4	461	12	BM712680	BM712680 UI-B-EJ0-
31	460	43.4	850	13	BU602918	BU602918 AGENCOURT
32	449	42.3	1280	13	BQ691500	BQ691500 AGENCOURT
C 33	443	41.8	446	12	BM681946	BM681946 UI-B-E01-
34	442	41.7	973	13	BU169156	BU169156 AGENCOURT
35	441	41.6	668	14	CB044866	CB044866 NISC_gc07
36	441	41.6	916	13	BQ954555	BQ954555 AGENCOURT
37	440	41.5	938	13	BU157842	BU157842 AGENCOURT
C 38	436	41.1	437	9	AI936228	AI936228 wq63e04.x
C 39	436	41.1	502	9	AI921394	AI921394 w024c07.x
40	423	39.9	780	29	AY419334	AY419334 Homo sapi
C 41	411	38.7	618	14	CD367193	CD367193 UI-H-FT2-
42	410	38.6	1127	13	BU174654	BU174654 AGENCOURT
C 43	402	37.9	1201	9	AL517150	AL517150 AL517150
C 44	399	37.6	451	9	AI493698	AI493698 qv97c07.x
C 45	392	36.9	874	13	BX362396	BX362396 BX362396

ALIGNMENTS

RESULT 1  
BQ575741/c  
LOCUS  
DEFINITION  
UI-H-EZ1-bbg-h-14-0-UI.s1 NCI CGAP Ch2 Homo sapiens cDNA clone  
UI-H-EZ1-bbg-h-14-0-UI 3', mRNA sequence.  
729 bp mRNA linear EST 19-JUN-2002  
ACCESSION  
BQ575741  
VERSION  
BQ575741.1 GI:21479058  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 729)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
Tumor Gene Index  
Unpublished (1997)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapsb@mail.nih.gov](mailto:cgapsb@mail.nih.gov)  
Tissue Procurement: Dr. Steven Gitelis/ Rush Presbyterian, Dept. of  
Orthopaedics  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
Seq primer: M13 FORWARD



Db 121 CAGGGGCGAGGAGAGAGATGCCCTGTCTCAGAGGATGCTGTGGCCCTGGAGAGC 180  
QY 287 ACAGTGTACAGCAACGAATCCAGAGCCGAGGTCTACGCCCGCCCTCGGCCACOGAC 346  
Db 181 ACAGTGTACAGCAACGAATCCAGAGCCGAGGTCTACGCCCGCCCTCGGCCACOGAC 240  
QY 347 CGCCTGGCCCTGCGCCCTTCCGCCAGCGGAGCGCTTCCACCGCTTCCAGCCCTAT 406  
Db 241 CGCCTGGCCCTGCGCCCTTCCGCCAGCGGAGCGCTTCCACCGCTTCCAGCCCTAT 300  
QY 407 CCGTACCTGCAGCAGAGATCGACCTGCGGCCACCATCTCGTGTTCAGACGGGAGGAG 466  
Db 301 CCGTACCTGCAGCAGAGATCGACCTGCGGCCACCATCTCGTGTTCAGACGGGAGGAG 360  
QY 467 CCCCCACCTACAGGGCCCTGACCCCTCGAGCTTCGGGACCCCGAGCAGCTGGAA 526  
Db 361 CCCCCACCTACAGGGCCCTGACCCCTCGAGCTTCGGGACCCCGAGCAGCTGGAA 420  
QY 527 CTGAACCGGGAGTCGGTGGCGCCACCCCAACAGAACCATCTTCGACAGTACCTGATG 586  
Db 421 CTGAACCGGGAGTCGGTGGCGCCACCCCAACAGAACCATCTTCGACAGTACCTGATG 480  
QY 587 GATAGTGCAGGCTGGCGGCCCTGCGGCCACCATCTTCGAGGAGTACGCGCCACG 646  
Db 481 GATAGTGCAGGCTGGCGGCCCTGCGGCCACCATCTTCGAGGAGTACGCGCCACG 540  
QY 647 TGCTACGGCAGCGGGCGGCGATGGAGGGCGCGCCACCATACAGCGAGGTCTATCGGC 706  
Db 541 TGCTACGGCAGCGGGCGGCGATGGAGGGCGCGCCACCATACAGCGAGGTCTATCGGC 600  
QY 707 CACTACCGGGTCTCTCTCCAGCACACAGAGCAGTGGCGCGCCCTCTCTGCTGGAG 766  
Db 601 CACTACCGGGTCTCTCTCCAGCACACAGAGCAGTGGCGCGCCCTCTCTGCTGGAG 660  
QY 767 GGGACCCGCTCCACCACACACATCGCGCCCTTAGAGAGCGGAGCCATCTGGAG 822  
Db 661 GGGACCCGCTCCACCACACACATCGCGCCCTTAGAGAGCGGAGCCATCTGGAG 716

RESULT 3  
BQ691705  
LOCUS  
DEFINITION  
AGENCOURT\_8046876 NIH\_MGC\_110 Homo sapiens cDNA clone IMAGE:6208939  
5', mRNA sequence.  
BQ691705  
BQ691705.1 GI:21817021  
KEYWORDS  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 655)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM2367 row: g column: 20  
High quality sequence stop: 645.  
Location/Qualifiers  
1. .655  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6208939"

/tissue type="ductal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_110"  
/note="Organ: pancreas; Vector: pOTB7; Site\_1: XhoI;  
Site\_2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 61.5%; Score 652; DB 13; Length 655;  
Best Local Similarity 100.0%; Pred. No. 2.le-154;  
Matches 652; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 410 TACCTGCAGCAGAGATCGACCTGCGGCCACCATCTCGTGTTCAGACGGGAGGAGCC 469  
Db 1 TACCTGCAGCAGAGATCGACCTGCGGCCACCATCTCGTGTTCAGACGGGAGGAGCC 60  
QY 470 CCACCCCTACAGGGCCCTTGCACCCCTCGAGCTTCGGGACCCCGAGCAGCTGGAATG 529  
Db 61 CCACCCCTACAGGGCCCTTGCACCCCTCGAGCTTCGGGACCCCGAGCAGCTGGAATG 120  
QY 530 AACCGGGAGTCGGTGGCGGCACCCCAACAGAACCATCTTCGACAGTACCTGATGGAT 589  
Db 121 AACCGGGAGTCGGTGGCGGCACCCCAACAGAACCATCTTCGACAGTACCTGATGGAT 180  
QY 590 AGTCCAGGCTGGCGGCCCTGCGGCCACCATCTTCGACAGTAACTCGGGCATCAGCCACGTGC 649  
Db 181 AGTCCAGGCTGGCGGCCCTGCGGCCACCATCTTCGACAGTAACTCGGGCATCAGCCACGTGC 240  
QY 650 TACGGCAGCGGGCGGCATGGAGGGCGCGCCACCATCTACAGCGAGGTCTATCGGACAC 709  
Db 241 TACGGCAGCGGGCGGCATGGAGGGCGCGCCACCATCTACAGCGAGGTCTATCGGACAC 300  
QY 710 TACCGGGGTCTCTCTTCAGCACACAGCAGCAGTGGCGCGCCCTCTCTGCTGGAGGGG 769  
Db 301 TACCGGGGTCTCTCTTCAGCACACAGCAGCAGTGGCGCGCCCTCTCTGCTGGAGGGG 360  
QY 770 ACCCGGCTCCACCACACACATCGCGGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAG 829  
Db 361 ACCCGGCTCCACCACACACATCGCGGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAG 420  
QY 830 AAGGATAACAGAAAGGACACCCCTCTTAGGGTCCCCAGGGGGCGCGGCTGGGCTGCG 889  
Db 421 AAGGATAACAGAAAGGACACCCCTCTTAGGGTCCCCAGGGGGCGCGGCTGGGCTGCG 480  
QY 890 TAGGTGAAAGGAGCAGAACACTCCGCGCTTCTTAGAAAGAGGAGTGAAGAGGGGGGG 949  
Db 481 TAGGTGAAAGGAGCAGAACACTCCGCGCTTCTTAGAAAGAGGAGTGAAGAGGGGGGG 540  
QY 950 CGCAGCAACGATCGTGTGGCCCTCCCTCCACCTCCCTGCTGTATATAATTTACATGT 1009  
Db 541 CGCAGCAACGATCGTGTGGCCCTCCCTCCACCTCCCTGCTGTATATAATTTACATGT 600  
QY 1010 GATGTCGCTCTGAATGCACAAGCTTAAGAGAGCTTGCAAAAAA 1061  
Db 601 GATGTCGCTCTGAATGCACAAGCTTAAGAGAGCTTGCAAAAAA 652

RESULT 4

AL558881/c  
LOCUS  
DEFINITION  
AL558881 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED  
Homo sapiens cDNA clone CS0D015YF12 3-PRIME, mRNA sequence.  
ACCESSION  
AL558881  
VERSION  
AL558881.2 GI:31283014  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
951 bp mRNA linear EST 31-MAY-2003





primer-adaptor  
[5'-pGACTAGTTCTAGATCGCGAGCGCGCC(T)15-3']. EST analysis  
was performed on the unamplified library at the NIH  
Intramural Sequencing Center (NISC).

ORIGIN

Query Match 57.4%; Score 609; DB 13; Length 609;  
Best Local Similarity 100.0%; Pred. No. 1.4e-143;  
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 214 CAGCCGGCAGCAGCCAGGGCGGAGGAGAGAGATGCCCTGTCTCAGAGGATGCCCTGTG 273  
|  
Db 1 CAGCCGGCAGCAGCCAGGGCGGAGGAGAGAGATGCCCTGTCTCAGAGGATGCCCTGTG 60  
|  
Qy 274 GCCCTCGAGAGACACAGTGTTCAGGCAACCGAATCCAGAGCCGCGAGTCTACGCCCGCC 333  
|  
Db 61 GCCCTCGAGAGACACAGTGTTCAGGCAACCGAATCCAGAGCCGCGAGTCTACGCCCGCC 120  
|  
Qy 334 TCGGCCACCGACCGCCTGGCCGTGCGGCCCTTCGCCAGCGGAGCGCTTCCACCGCTT 393  
|  
Db 121 TCGGCCACCGACCGCCTGGCCGTGCGGCCCTTCGCCAGCGGAGCGCTTCCACCGCTT 180  
|  
Qy 394 CCAGCCACCTATCCGTACCTGACGACGAGATCGACCTGCCGCCACCATCTCGTGTG 453  
|  
Db 181 CCAGCCACCTATCCGTACCTGACGACGAGATCGACCTGCCGCCACCATCTCGTGTG 240  
|  
Qy 454 AGAGGGGAGGAGCCCCACCCCTACCAAGGCGCCCTGACCCCTCCAGCTTCGGGACCCCGA 513  
|  
Db 241 AGAGGGGAGGAGCCCCACCCCTACCAAGGCGCCCTGACCCCTCCAGCTTCGGGACCCCGA 300  
|  
Qy 514 GCAACGAGCTGGAACCTGAACCGGAGTGGGTGGCGCACCCCAACAGAACCATCTTCCA 573  
|  
Db 301 GCAGGAGCTGGAACCTGAACCGGAGTGGGTGGCGCACCCCAACAGAACCATCTTCCA 360  
|  
Qy 574 CAGTGACCTGATGATAGTGCAGGCTGGGGCGGCCCTGCGGCCCTCCAGCAGTAACCTCGGG 633  
|  
Db 361 CAGTGACCTGATGATAGTGCAGGCTGGGGCGGCCCTGCGGCCCTCCAGCAGTAACCTCGGG 420  
|  
Qy 634 CATCAGGCCACGTGCTACGGCAGCGGCGCGCATGGAGGGCGCGCATGGAGGGCGCGCATACAG 693  
|  
Db 421 CATCAGGCCACGTGCTACGGCAGCGGCGCGCATGGAGGGCGCGCATGGAGGGCGCGCATACAG 480  
|  
Qy 694 CGAGTCTATCGGCCACTACCGGGGTCTCTCTCCAGCACACACATCGCGCCCTAGAGAGCGCAGC 753  
|  
Db 481 CGAGTCTATCGGCCACTACCGGGGTCTCTCTCCAGCACACACATCGCGCCCTAGAGAGCGCAGC 540  
|  
Qy 754 CTCCTGTCTGGAGGGGACCGGCTCCACACACACATCGCGCCCTAGAGAGCGCAGC 813  
|  
Db 541 CTCCTGTCTGGAGGGGACCGGCTCCACACACACATCGCGCCCTAGAGAGCGCAGC 600  
|  
Qy 814 CATCTGGAG 822  
|  
Db 601 CATCTGGAG 609  
|

RESULT 6

BQ690750  
LOCUS BQ690750  
DEFINITION AGENCOURT\_8046394 NIH\_MGC\_110 Homo sapiens cDNA clone IMAGE:6209341  
5', mRNA sequence.

ACCESSION BQ690750

VERSION BQ690750.1 GI:21816066

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 890)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

CONTACT Robert Strausberg, Ph.D.

EMAIL: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LCM2368 row: h column: 14  
High quality sequence stop: 627.

FEATURES  
source

1..890  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6209341"  
/tissue\_type="ductal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_110"  
/note="Organ: pancreas; Vector: pOTB7; Site:1: XhoI;  
Site:2: EcoRI; CDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 56.6%; Score 601; DB 13; Length 890;  
Best Local Similarity 99.8%; Pred. No. 1e-141;  
Matches 651; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TCCTCCTTGGGTTGGGTGAAGCGCTTCGGGGTTTCAGTGGGCCATGATCCCCGAGCTGC 60  
|  
Db 24 TCCTCCTTGGGTTGGGTGAAGCGCTTCGGGGTTTCAGTGGGCCATGATCCCCGAGCTGC 83  
|  
Qy 61 TGGAGAACTGAAGCGCGGACGGTCTCTCGCGAAACAGGCAATGGCGGAGTGGAGTTGT 120  
|  
Db 84 TGGAGAACTGAAGCGCGGACGGTCTCTCGCGAAACAGGCAATGGCGGAGTGGAGTTGT 143  
|  
Qy 121 TCAGATCATCATCATCGTGGTGGTGAATGATGTTGTTGTTGATCACTGCTGCT 180  
|  
Db 144 TCAGATCATCATCATCGTGGTGGTGAATGATGTTGTTGTTGATCACTGCTGCT 203  
|  
Qy 181 GAGCCACTACAAGCTGTCTGACGGTCTCTTCATCAGCCGGCACAGCCAGGCGGAGGAG 240  
|  
Db 204 GAGCCACTACAAGCTGTCTGACGGTCTCTTCATCAGCCGGCACAGCCAGGCGGAGGAG 263  
|  
Qy 241 AGAGATGCCCTGTCTCTCAGAGGATGCTCTGTGGCCCTCGGAGAGCACAGTGTTCAGGCAA 300  
|  
Db 264 AGAGATGCCCTGTCTCTCAGAGGATGCTCTGTGGCCCTCGGAGAGCACAGTGTTCAGGCAA 323  
|  
Qy 301 CGGAATCCCAGAGCCCGCAGGTCTACGCCCGCCCTCGGCCCCACCGACCGCTGGCCGTGCC 360  
|  
Db 324 CGGAATCCCAGAGCCCGCAGGTCTACGCCCGCCCTCGGCCCCACCGACCGCTGGCCGTGCC 383  
|  
Qy 361 GCCCTTGGCCCGAGCGGAGCGGCTTCCACCGCTTCCAGCCCACTATCGGTACCTGCAGCA 420  
|  
Db 384 GCCCTTGGCCCGAGCGGAGCGGCTTCCACCGCTTCCAGCCCACTATCGGTACCTGCAGCA 443  
|  
Qy 421 CGAGATCGACCTGCCCGCCACCATCTCGCTGTTCAGACGGGAGGAGGCCCCACCTACCA 480  
|  
Db 444 CGAGATCGACCTGCCCGCCACCATCTCGCTGTTCAGACGGGAGGAGGCCCCACCTACCA 503  
|  
Qy 481 GGGCCCTGACCCCTCCAGCTTCGGGACCCCGAGCAGCTGGAACTGAACCGGGAGTC 540  
|  
Db 504 GGGCCCTGACCCCTCCAGCTTCGGGACCCCGAGCAGCTGGAACTGAACCGGGAGTC 563  
|  
Qy 541 GGTGCGGCGACCCCGCAACAGAACCATCTTCGACAGTGCCTGTATGATGATGCCAGGCT 600  
|  
Db 564 GGTGCGGCGACCCCGCAACAGAACCATCTTCGACAGTGCCTGTATGATGATGCCAGGCT 623  
|  
Qy 601 GGGCGGCGCCCTGCGCCCGCCAGCAGTAACCTCGGGCATCAGCGCCACCTGTGCTAC 652  
|



```
|||||
Db      624 GGGGGCCCTGCCCCCCCCAGCAGTAACCTCGGCATCAGCGCCAGTGCTAC 675

RESULT 7
CA431191/c
LOCUS
DEFINITION
  UI-H-FG1-bgi-c-12-0-UI.s1 602 bp mRNA linear EST 07-NOV-2002
  UI-H-FG1-bgi-c-12-0-UI 3', mRNA sequence.
CA431191
ACCESSION
VERSION
KEYWORDS
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 602)
AUTHORS
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
JOURNAL
COMMENT
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-re@mail.nih.gov
  Tissue Procurement: James Martin
  cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Clone distribution information can be obtained
  from Dr. M. Bento Soares, bento-soares@uiowa.edu
  Seq primer: M13 FORWARD
  POLYA=Yes.
  Location/Qualifiers
    1..602
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="UI-H-FG1-bgi-c-12-0-UI"
      /tissue_type="Cell lines"
      /dev_stage="Adult"
      /lab_host="DH10B (Life Technologies)"
      /clone_lib="NCI CGAP FG1"
      /note="Organ: Enchondroma; Vector: pT7T3-Pac (Pharmacia)
      with a modified polylinker; Site 1: Ecor I; Site 2: Not I;
      NCI CGAP_FG1 is a normalized cDNA library obtained from a
      pool of mRNA from 2 cell lines from Enchondroma tissues.
      The library was constructed according to Bonaldo, Lennon
      and Soares, Genome Research, 6:791-806, 1996. First strand
      cDNA synthesis was primed with an oligo-dT primer
      containing a Not I site. Double stranded cDNA was ligated
      to an Ecor I adaptor, digested with Not I, and cloned
      directionally into pT7T3-Pac vector. The oligonucleotide
      used to prime the synthesis of first-strand cDNA contains
      a library tag sequence that is located between the Not I
      site and the (dT)18 tail. The sequence tag for this
      library is CCGTCACTC. The cell lines were provided by Dr.
      James Martin from the University of Iowa.
      TAG TISSUE=Enchondroma cell line (Mix of EN1 and EN2)
      TAG_LIB=UI-H-FG1
      TAG_SEQ=CGGTCACTC

FEATURES
  source
    1..602
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="UI-H-FG1-bgi-c-12-0-UI"
      /tissue_type="Cell lines"
      /dev_stage="Adult"
      /lab_host="DH10B (Life Technologies)"
      /clone_lib="NCI CGAP FG1"
      /note="Organ: Enchondroma; Vector: pT7T3-Pac (Pharmacia)
      with a modified polylinker; Site 1: Ecor I; Site 2: Not I;
      NCI CGAP_FG1 is a normalized cDNA library obtained from a
      pool of mRNA from 2 cell lines from Enchondroma tissues.
      The library was constructed according to Bonaldo, Lennon
      and Soares, Genome Research, 6:791-806, 1996. First strand
      cDNA synthesis was primed with an oligo-dT primer
      containing a Not I site. Double stranded cDNA was ligated
      to an Ecor I adaptor, digested with Not I, and cloned
      directionally into pT7T3-Pac vector. The oligonucleotide
      used to prime the synthesis of first-strand cDNA contains
      a library tag sequence that is located between the Not I
      site and the (dT)18 tail. The sequence tag for this
      library is CCGTCACTC. The cell lines were provided by Dr.
      James Martin from the University of Iowa.
      TAG TISSUE=Enchondroma cell line (Mix of EN1 and EN2)
      TAG_LIB=UI-H-FG1
      TAG_SEQ=CGGTCACTC

ORIGIN
  Query Match 56.5%; Score 599; DB 14; Length 602;
  Best Local Similarity 100.0%; Pred.No. 4.5e-141;
  Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      463 GGAGCCCCCACCCTACAGGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCT 522
Db      602 GGAGCCCCCACCCTACAGGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCT 543
QY      523 GGAACCTGAACCGGGAGTCGGTGGCGGACCCGCCCAACAGAACCACTTCGACAGTGACCT 582
Db      542 GGAACCTGAACCGGGAGTCGGTGGCGGACCCGCCCAACAGAACCACTTCGACAGTGACCT 483

QY      583 GATGATAGTCCAGGCTGGGGGGCCCCCTGCCCCCCCCAGCAGTAACCTCGGCATCAGCGC 642
Db      482 GATGATAGTCCAGGCTGGGGGGCCCCCTGCCCCCCCCAGCAGTAACCTCGGCATCAGCGC 423
QY      643 CACGTGCTACCGGCAGCGCGGGCGCATGGAGGGGGCGCGCCGCCACCTACAGCGAGGTCT 702
Db      422 CACGTGCTACCGGCAGCGCGGGCGCATGGAGGGGGCGCGCCGCCACCTACAGCGAGGTCT 363
QY      703 CGGCCACTACCGGGGTCTCTCTTCCAGCACCACTAGGAGCAGTGGGGCGCCCTCTCTTCT 762
Db      362 CGGCCACTACCGGGGTCTCTCTTCCAGCACCACTAGGAGCAGTGGGGCGCCCTCTCTTCT 303
QY      763 GGAGGGGACCCGGCTCCACCACACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAG 822
Db      302 GGAGGGGACCCGGCTCCACCACACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAG 243
QY      823 CAAGAGAGGATAAACAGAAAGGACACCCCTCTCTAGGGTCCCGAGGGGGCGGGGTGG 882
Db      242 CAAGAGAGGATAAACAGAAAGGACACCCCTCTCTAGGGTCCCGAGGGGGCGGGGTGG 183
QY      883 GGCTGCTAGGTGAAAAGGCGAGAACACTCCGGCTCTCTTAGAGAGGAGTGAGAGGAGG 942
Db      182 GGCTGCTAGGTGAAAAGGCGAGAACACTCCGGCTCTCTTAGAGAGGAGTGAGAGGAGG 123
QY      943 CGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATT 1002
Db      122 CGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATT 63
QY      1003 TACATGTGATGTCTGTCTGAATGCACAAAGCTAAGAGAGCTTGCAAAAAA 1061
Db      62 TACATGTGATGTCTGTCTGAATGCACAAAGCTAAGAGAGCTTGCAAAAAA 4

RESULT 8
BU624784/c
LOCUS
DEFINITION
  UI-H-FG1-bgi-f-22-0-UI.s1 588 bp mRNA linear EST 23-SEP-2002
  UI-H-FG1-bgi-f-22-0-UI 3', mRNA sequence.
BU624784
ACCESSION
VERSION
KEYWORDS
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 588)
AUTHORS
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
JOURNAL
COMMENT
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-re@mail.nih.gov
  Tissue Procurement: James Martin
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Clone distribution information can be obtained
  from Dr. M. Bento Soares, bento-soares@uiowa.edu
  Seq primer: M13 FORWARD
  POLYA=Yes.
  Location/Qualifiers
    1..588
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="UI-H-FG1-bgi-f-22-0-UI"
      /tissue_type="Cell lines"
      /dev_stage="Adult"
      /lab_host="DH10B (Life Technologies)"
      /clone_lib="NCI CGAP FG1"
      /note="Organ: Enchondroma; Vector: pT7T3-Pac (Pharmacia)
      with a modified polylinker; Site 1: Ecor I; Site 2: Not I;
      NCI CGAP_FG1 is a normalized cDNA library obtained from a
```





QY 728 CAGCACCAGCAGAGAGTGGCGCGCCCTCTTCTGCTGAGGGGACCGGCTCCACACACA 787  
Db |||||||  
QY 788 CACATCGCGCCCTAGAGAGCGCAGCCCATCTGGAGCAAGAGAGATAAACAGAAAGGA 847  
Db |||||||  
QY 277 CACATCGCGCCCTAGAGAGCGCAGCCCATCTGGAGCAAGAGAGATAAACAGAAAGGA 218  
Db |||||||  
QY 848 CACCCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGGCTGGCTAGGTGAAAGGCGAGAAC 907  
Db |||||||  
QY 217 CACCCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGGCTGGCTAGGTGAAAGGCGAGAAC 158  
Db |||||||  
QY 908 ACTCCGCGCTCTTAGAGAGGAGTGAGAGGAAGGCGGGGGCGGCGAGCAACGATCGTGT 967  
Db |||||||  
QY 157 ACTCCGCGCTCTTAGAGAGGAGTGAGAGGAAGGCGGGGGCGGCGAGCAACGATCGTGT 98  
Db |||||||  
QY 968 GGCCCTCCCTCCACCTCCCTGTGTATATAATATTATCATGTGATGTCTGCTGAATGC 1027  
Db |||||||  
QY 97 GGCCCTCCCTCCACCTCCCTGTGTATATAATATTATCATGTGATGTCTGCTGAATGC 38  
Db |||||||  
QY 1028 ACAAGCTAAGAGAGCTTGCACAAAAA 1057  
Db |||||||  
QY 37 ACAAGCTAAGAGAGCTTGCACAAAAA 8  
Db |||||||

RESULT 10  
BE855409/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BB855409 559 bp mRNA linear EST 29-SEP-2000  
7913f05.x1 NCI CGAP Brn23 Homo sapiens cDNA clone IMAGE:3306369 3'  
similar to TR:Q9UJD3 Q9UJD3 DJ718J7.1 ;, mRNA sequence.

BB855409  
BE855409.1 GI:10367404  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 559)  
NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute / National Institute of Neurological  
Disorders and Stroke, Brain Tumor Genome Anatomy Project  
(CGAP/BTGP), Tumor Gene Index  
Unpublished (1998)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,  
Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonardo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL, send email to:  
info@image.llnl.gov  
Seq primer: -40UP from Gibco  
High quality sequence stop: 443.  
Location/Qualifiers  
1. .559  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3306369"  
/tissue\_type="glioblastoma (pooled)"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Brn23"  
/note="Organ: brain; Vector: p7T3D-Pac (Pharmacia) with a  
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer [5'  
TGTTACCAATCTGAAGTGGGAGCGCGCGATATCTTTTTTTTTTTTTTTTTTTT  
T 3']; double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified p7T3 vector.

Library is normalized, and was constructed by Bento  
Soares and M.Fatima Bonardo."

ORIGIN  
Query Match 52.7%; Score 559; DB 10; Length 559;  
Best Local Similarity 100.0%; Pred. No. 5.2e-131;  
Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 488 TGCACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACTGAACCGGGAGTCGGTGGC 547  
Db |||||||  
QY 559 TGCACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACTGAACCGGGAGTCGGTGGC 500  
Db |||||||  
QY 548 GCACCCCAACAGAACCATCTTCGACAGTACCTGATGATAGTCCAGGCTGGGGGC 607  
Db |||||||  
QY 499 GCACCCCAACAGAACCATCTTCGACAGTACCTGATGATAGTCCAGGCTGGGGGC 440  
Db |||||||  
QY 608 CCCTGCCCCCAGCACTAACTCGGGCATCAGCGCCACCTGCTACGGCAGCGGGGC 667  
Db |||||||  
QY 439 CCCTGCCCCCAGCACTAACTCGGGCATCAGCGCCACCTGCTACGGCAGCGGGGC 380  
Db |||||||  
QY 668 ATGGAGGGGGCGCGCCACCTACAGCGAGGTCTCGCCACTACCGGGCTCCCTTC 727  
Db |||||||  
QY 379 ATGGAGGGGGCGCGCCACCTACAGCGAGGTCTCGCCACTACCGGGCTCCCTTC 320  
Db |||||||  
QY 728 CAGCAACAGCAGAGCAGTGGGGCGCGCCCTCTTCTGAGGGGACCGGCTCCACACACA 787  
Db |||||||  
QY 319 CAGCAACAGCAGAGCAGTGGGGCGCGCCCTCTTCTGAGGGGACCGGCTCCACACACA 260  
Db |||||||  
QY 788 CACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGATAAACAGAAAGGA 847  
Db |||||||  
QY 259 CACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGATAAACAGAAAGGA 200  
Db |||||||  
QY 848 CACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGGCTGGGTGAGGTGAAAGGCGAGAAC 907  
Db |||||||  
QY 199 CACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGGCTGGGTGAGGTGAAAGGCGAGAAC 140  
Db |||||||  
QY 908 ACTCGCGCTCTTAGAGAGGAGTGAGAGGAAGGCGGGGGCGGCGAGCAACGATCGTGT 967  
Db |||||||  
QY 139 ACTCGCGCTCTTAGAGAGGAGTGAGAGGAAGGCGGGGGCGGCGAGCAACGATCGTGT 80  
Db |||||||  
QY 968 GGCCCTCCCTCCACCTCCCTGTGTATAATAATATTATCATGTGATGTCTGCTGAATGC 1027  
Db |||||||  
QY 79 GGCCCTCCCTCCACCTCCCTGTGTATAATAATATTATCATGTGATGTCTGCTGAATGC 20  
Db |||||||  
QY 1028 ACAAGCTAAGAGAGCTTGC 1046  
Db |||||||  
QY 19 ACAAGCTAAGAGAGCTTGC 1

RESULT 11  
BM974296/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

BM974296 626 bp mRNA linear EST 20-FEB-2003  
UI-CF-EC1-aca-k-23-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone  
UI-CF-EC1-aca-k-23-0-UI 3', mRNA sequence.  
BM974296  
BM974296.1 GI:19591887  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 626)  
Bonardo, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
8889548  
Contact: McCray, PB  
McCray Lab  
University of Iowa  
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
Tel: 319 356 4866



Fax: 319 356 7171  
 Email: paul-mccray@uiowa.edu  
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com) or from Open Biosystems (www.openbiosystems.com).  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

FEATURES  
 source  
 1. .626  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-CF-EC1-aca-k-23-0-UI"  
 /tissue\_type="Lung"  
 /dev\_stage="Adult and Fetal"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-CF-EC1"  
 /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-EC1 is a normalized cDNA library containing the following tissue(s): Normal lung from adult and from fetal day 64, day 87, week 19 and week 42. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AAGTCTTAC.  
 TAG TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371 and 380-383  
 TAG LIB=UI-CF-EC1  
 TAG\_SEQ=AAGTCTTAC"

ORIGIN

Query Match 52.4%; Score 556; DB 12; Length 626;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-130;  
 Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 506 GACCCGAGCAGCAGCTGGAACCTGAACCGGAGTGGTGGCGGACCCCAACAGAAC 565  
 Db |||||  
 QY 559 GACCCGAGCAGCAGCTGGAACCTGAACCGGAGTGGTGGCGGACCCCAACAGAAC 500  
 Db |||||  
 QY 566 ATCTTCGACAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 625  
 Db |||||  
 QY 499 ATCTTCGACAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 440  
 Db |||||  
 QY 626 AACTCGGGCATCAGCGCCACGTGCTACGGCAGCGCGGGCGGCAATGAGGCGCGCC 685  
 Db |||||  
 QY 439 AACTCGGGCATCAGCGCCACGTGCTACGGCAGCGCGGGCGGCAATGAGGCGCGCC 380  
 Db |||||  
 QY 686 ACCTACAGCGAGTTCATCGGCCACTACCGCGGTCTCTTCCAGCACACAGAGCAGT 745  
 Db |||||  
 QY 379 ACCTACAGCGAGTTCATCGGCCACTACCGCGGTCTCTTCCAGCACACAGAGCAGT 320  
 Db |||||  
 QY 746 GGGCGGCCCTCTTGTGGAGGGGACCGGCTCCACCACACACATCGCGCCCTAGAG 805  
 Db |||||  
 QY 319 GGGCGGCCCTCTTGTGGAGGGGACCGGCTCCACCACACACATCGCGCCCTAGAG 260  
 Db |||||  
 QY 806 AGCGCAGCCATCTGGAGCAAGAGAGAGGATAAACAGAAAGGACACCTCTTAGGGTCC 865  
 Db |||||  
 QY 259 AGCGCAGCCATCTGGAGCAAGAGAGAGGATAAACAGAAAGGACACCTCTTAGGGTCC 200  
 Db |||||  
 QY 866 CAGGGGGCGGGCTGGGGCTGCGTAGGTGTAAGAGGAGAGAACTCCGGCTTCTTAGAA 925  
 Db |||||  
 QY 199 CAGGGGGCGGGCTGGGGCTGCGTAGGTGTAAGAGGAGAGAACTCCGGCTTCTTAGAA 140  
 Db |||||

QY 926 GAGGAGTGAGAGAGCGGG 985  
 Db |||||  
 QY 139 GAGGAGTGAGAGAGCGGG 80  
 Db |||||  
 QY 986 CCCTGTGTATAATATTACATGTGATGTCTGTCTGTGAATGCACAAGCTAAGAGAGCTTG 1045  
 Db |||||  
 QY 79 CCCTGTGTATAATATTACATGTGATGTCTGTCTGTGAATGCACAAGCTAAGAGAGCTTG 20  
 Db |||||  
 QY 1046 CAAAAAATAAAAAA 1061  
 Db |||||  
 QY 19 CAAAAAATAAAAAA 4

RESULT 12  
 BQ575582/c  
 LOCUS  
 DEFINITION  
 BQ575582  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 570)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: Dr. Steven Gitelis/ Rush Presbyterian, Dept. of Orthopaedics  
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

FEATURES  
 source  
 1. .570  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-H-EZ1-bbf-f-09-0-UI"  
 /tissue\_type="Chondrosarcoma Grade II"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI CGAP Ch2"  
 /note="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Ch2 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma Grade II. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCT.  
 TAG TISSUE=grade-2-chondrosarcoma  
 TAG LIB=UI-H-EZ1  
 TAG\_SEQ=ATCTAATATG"

ORIGIN

Query Match 50.5%; Score 536; DB 13; Length 570;  
 Best Local Similarity 100.0%; Pred. No. 3e-125;



```
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 522 TGAAGTGAACCGGAGTCGGTGCAGCACCACCCAAACAGAACCATCTTCGACAGTGACC 581
Db 543 TGAAGTGAACCGGAGTCGGTGCAGCACCACCCAAACAGAACCATCTTCGACAGTGACC 484
QY 582 TGATGGATAGTCCAGGTCGGGGCGCCCTCTCCAGCAGTAACTCGGGGATCAGCG 641
Db 483 TGATGGATAGTCCAGGTCGGGGCGCCCTCTCCAGCAGTAACTCGGGGATCAGCG 424
QY 642 CCACGTGTACCGGAGCGGGCGGCATGGAGGGCGCGCCACCTACAGGAGGTCA 701
Db 423 CCACGTGTACCGGAGCGGGCGGCATGGAGGGCGCGCCACCTACAGGAGGTCA 364
QY 702 TCGGCCACTACCGGGGTCTCTCTCCAGCAGCAGCAGTGGGGCGCCCTCTCTTGC 761
Db 363 TCGGCCACTACCGGGGTCTCTCTCCAGCAGCAGCAGTGGGGCGCCCTCTCTTGC 304
QY 762 TGGAGGGGACCGGGTCTCCACACACACATCGGGCCCTAGAGAGCGCAGCCATCTGA 821
Db 303 TGGAGGGGACCGGGTCTCCACACACATCGGGCCCTAGAGAGCGCAGCCATCTGA 244
QY 822 GCAAGAGAGAGTAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTG 881
Db 243 GCAAGAGAGAGTAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTG 184
QY 882 GGGCTGCGTAGTGAAGAGCAGAACACTCCGGCTCTTAGAAGAGGAGTGAGAGGAAG 941
Db 183 GGGCTGCGTAGTGAAGAGCAGAACACTCCGGCTCTTAGAAGAGGAGTGAGAGGAAG 124
QY 942 GCGGGGGCGCAGCAGCATCGTGTGGCCCTCCCTCCAGCTCCCTCTCTGTATATAAT 1001
Db 123 GCGGGGGCGCAGCAGCATCGTGTGGCCCTCCCTCCAGCTCCCTCTCTGTATATAAT 64
QY 1002 TTACATGTGATGTCTGTCTGAATGCACAGCTTAAGAGAGCTTGCAAAAAA 1057
Db 63 TTACATGTGATGTCTGTCTGAATGCACAGCTTAAGAGAGCTTGCAAAAAA 8
```

RESULT 13  
BU539219  
LOCUS  
DEFINITION  
AGENCOURT 10215265 NIH\_MGC\_107 Homo sapiens cDNA clone  
IMAGE:6569922 5', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BU539219  
BU539219.1 GI:22849660  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 945)  
NIH-MGC <http://mgi.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LLM2757 row: p column: 18  
High quality sequence stop: 663.  
Location/Qualifiers

FEATURES  
source  
1. .945  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6569922"  
/tissue\_type="adenocarcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_107"  
/note="Organ: breast; Vector: pOTB7; Site\_1: EcoRI;  
Site\_2: XhoI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 50.5%; Score 536; DB 13; Length 945;  
Best Local Similarity 99.5%; Pred. No. 2e-125;  
Matches 736; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 18 TGAAGGCGCTTGGGGTTTTCAGTGGGCCATGATCCCGAGCTGCTGGAGAACTGAAGGCGG 77  
Db 1 TGAAGGCGCTTGGGGTTTTCAGTGGGCCATGATCCCGAGCTGCTGGAGAACTGAAGGCGG 60  
QY 78 ACGGTCTCTCGGAACACAGGCAATGGGGAGCTGGAGTTTGTTCAGATCATCATCG 137  
Db 61 ACAGTCTCTCGGAACACAGGCAATGGGGAGCTGGAGTTTGTTCAGATCATCATCG 120  
QY 138 TGGTGGTGTATGATGGTGGTGGTGTGATCACGTGCTGCTGAGCCACTACAAGCTGT 197  
Db 121 TGGTGGTGTATGATGGTGGTGGTGTGATCACGTGCTGCTGAGCCACTACAAGCTGT 180  
QY 198 CTGCACGGTCTTTCATCAGCCGGCAGCCAGGGGGGAGGAGAGATGCCCTGTCT 257  
Db 181 CTGCACGGTCTTTCATCAGCTGGCAGCAGCCAGGGGGGAGGAGAGATGCCCTGTCT 240  
QY 258 CAGAAGGATGCTGTGGCCCTCGGAGAGCACAGTGTTCAGGCAACGGGAATCCAGAGCGC 317  
Db 241 CAGAAGGATGCTGTGGCCCTCGGAGAGCACAGTGTTCAGGCAACGGGAATCCAGAGCGC 300  
QY 318 AGGTCTACGCCCCCGCTCGGCCCCACCGACCGCTGCGCCCTTCCGCGCCAGCGG 377  
Db 301 AGGTCTACGCCCCCGCTCGGCCCCACCGACCGCTGCGCCCTTCCGCGCCAGCGG 360  
QY 378 AGCGCTTCCACCGCTTCCAGCCCCACCTATCCGTACTTCAGCAGCAGATGCGCCG 437  
Db 361 AGCGCTTCCACCGCTTCCAGCCCCACCTATCCGTACTTCAGCAGCAGATGCGCCG 420  
QY 438 CCACCATCTCGTGTTCAGACGGGGAGGAGCCCCCACCCTACAGGGCCCCCTGCCACCTCC 497  
Db 421 CCACCATCTCGTGTTCGGAAGGGGAGGAGCCCCCACCCTACAGGGCCCCCTGCCACCTCC 480  
QY 498 AGCTTCGGGACCCCGAGCAGCAGCTGGAACTGAAACCGGGAGTGGTGGCGGCAACCCCA 557  
Db 481 AGCTTCGGGACCCCGAGCAGCAGCTGGAACTGAAACCGGGAGTGGTGGCGGCAACCCCA 540  
QY 558 ACAGAACCATCTTCGACAGTGAACCTGATGGATAGTCCAGGCTGGGGGGCCCCCTGCCCCC 617  
Db 541 ACAGAACCATCTTCGACAGTGAACCTGATGGATAGTCCAGGCTGGGGGGCCCCCTGCCCCC 600  
QY 618 CCAGCAGTAACTCGGGCATCAGCGCCACGTGTACGGCAGCGGGCGGCGCATGGAGGGGC 677  
Db 601 CCAGCAGTAACTCGGGCATCAGCGCCACGTGTACGGCAGCGGGCGGCGCATGGAGGGGC 660  
QY 678 GCGCGCCACCTACAGGAGGTTCATCGGCCACTACCGGGGTCTCTCTTCCAGACCAAGC 737  
Db 661 GCGCGCCACCTACAGGAGGTTCATCGGCCACTACCGGGGTCTCTCTTCCAGACCAAGC 720  
QY 738 AGAGCAGTGGGGCGCCCTCC 757  
Db 721 AGAGCAGTGGGGCGCCCTCC 740

RESULT 14  
CB049800  
LOCUS  
DEFINITION  
NISC\_gj13c01.y1 NCI\_CGAP\_Pr28 Homo sapiens cDNA clone IMAGE:3271656

5', mRNA sequence.  
 CB049800  
 CB049800.1 GI:27788087  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 563)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
 cDNA Library Preparation:  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC)  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[info@image.llnl.gov](mailto:info@image.llnl.gov)  
 Plate: LLAM8009 row: E column: 1  
 Seq primer: M13RPI reverse primer (ABI).  
 Location/Qualifiers  
 1..563  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3271656"  
 /sex="male"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP Pr28"  
 /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)  
 with a modified polylinker; Plasmid DNA from the  
 normalized library NCI CGAP Pr22 was prepared, and ss  
 circles were made in vitro. Following HAP purification,  
 this DNA was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from a pool  
 of 5,000 clones made from the same library (cloneIDs  
 985608-986759, 1101192-1101959, and 1217928-1220615).  
 Subtraction by Bento Soares and M. Fatima Bonaldo."

Accession  
 Version  
 Keywords  
 Source  
 Organism  
 Reference  
 Authors  
 Title  
 Journal  
 Comment  
 Features  
 source  
 Origin  
 Query Match 50.0%; Score 530; DB 14; Length 563;  
 Best Local Similarity 100.0%; Pred. No. 9.7e-124;  
 Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 528 TGAACCGGGAGTCGGTGGCGGACCCCAACAGAACCATCTTCGACAGTGCACCTGATGG 587  
 10 TGAACCGGGAGTCGGTGGCGGACCCCAACAGAACCATCTTCGACAGTGCACCTGATGG 69  
 588 ATAGTGCCAGGCTGGCGGCGCCCTGCGCCCGCCAGCAGTAACCTGGGCAATCAGCGCCACGT 647  
 70 ATAGTGCCAGGCTGGCGGCGCCCTGCGCCCGCCAGCAGTAACCTGGGCAATCAGCGCCACGT 129  
 648 GCTACGGCAGCGCGGGCGCATGGAGGGGCGCGCCCGCCACCTACAGCGAGTGCATCGGCC 707  
 130 GCTACGGCAGCGCGGGCGCATGGAGGGGCGCGCCCGCCACCTACAGCGAGTGCATCGGCC 189  
 708 ACTACCGGGTCTCTCTCCAGCACCAGCAGTGGCGCGCCCTCTCTGCTGGAGG 767  
 190 ACTACCGGGTCTCTCTCCAGCACCAGCAGTGGCGCGCCCTCTCTGCTGGAGG 249  
 768 GGACCGGGTCTCACCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAAG 827  
 250 GGACCGGGTCTCACCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAAG 309  
 828 AGAAGGATAACAGAAAGGACACCCCTCTCTAGGGTCCCAGGGGGCGCGGGCTGGGGCTG 887  
 310 AGAAGGATAACAGAAAGGACACCCCTCTCTAGGGTCCCAGGGGGCGCGGGCTGGGGCTG 369

888 CGTAGGTGAAAGGCGAGAACACTCCGGCTTCTTAGAAGAGAGTGAAGAGCGCGGG 947  
 370 CGTAGGTGAAAGGCGAGAACACTCCGGCTTCTTAGAAGAGAGTGAAGAGCGCGGG 429  
 948 GCGCGAGCAACGCATCGTGGCCCTCCCTCCACCTCCCTCCCTCTGTGTATTAATTTACAT 1007  
 430 GCGCGAGCAACGCATCGTGGCCCTCCCTCCACCTCCCTCCCTCTGTGTATTAATTTACAT 489  
 1008 GTGATGTCTGGTCTGATGACCAAGCTAAGAGAGCTTGCAGAGAGAGAGAGAGAGAG 1057  
 490 GTGATGTCTGGTCTGATGACCAAGCTAAGAGAGCTTGCAGAGAGAGAGAGAGAGAG 539

Result 15  
 BQ686793  
 LOCUS  
 DEFINITION  
 AGENCOURT\_8345390 NIH\_MGC\_110 Homo sapiens cDNA clone IMAGE:6248044  
 5', mRNA sequence.  
 BQ686793  
 BQ686793.1 GI:21812109  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 844)  
 NIH-MGC <http://mgi.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LCM2387 row: e column: 05  
 High quality sequence stop: 593.  
 Location/Qualifiers  
 1..844  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6248044"  
 /tissue\_type="ductal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_110"  
 /note="Organ: pancreas; Vector: pOTB7; Site: 1: XhoI;  
 Site 2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

Query Match 49.8%; Score 528; DB 13; Length 844;  
 Best Local Similarity 99.8%; Pred. No. 2.3e-123;  
 Matches 578; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 483 GCCCTGACCCCTCCAGCTTCGGGACCCCGAGCAGCTGGAACCTGAACCGGAGTCGG 542  
 1 GCCCTGACCCCTCCAGCTTCGGGACCCCGAGCAGCTGGAACCTGAACCGGAGTCGG 60  
 543 TGCAGCGACCCCAACAGAACCATCTTCGACAGTGCACCTGATGATGATGATGATGATG 602  
 61 TGCAGCGACCCCAACAGAACCATCTTCGACAGTGCACCTGATGATGATGATGATGATG 120  
 603 GCGGCCCTGCCCCCAGCAGTAACCTCGGGCATCAGCGCCACCTGCTACCGCAGCGCG 662

Db 121 GCGGCCCTGCCCCCAGCAGTAACCTGGGCATCAGGCCACGTGCTACGGCAGCGGCG 180  
Qy 663 GGCGCATGGAGGGCGCGCCACCTACAGCGAGGTCACTGGGCCACTACCCGGGGTCCT 722  
Db 181 GGCGCATGGAGGGCGCGCCACCTACAGCGAGGTCACTGGGCCACTACCCGGGGTCCT 240  
Qy 723 CCTTCCAGCACCCAGCAGAGCAGTGGGCGGCCCTCTTGCTGAGGGGACCCGGCTCCACC 782  
Db 241 CCTTCCAGCACCCAGCAGAGCAGTGGGCGGCCCTCTTGCTGAGGGGACCCGGCTCCACC 300  
Qy 783 ACACACACATCGCGCCCTTAGAGAGCGCAGCCCATCTGAGCAAGAGAAAGTAAACAGA 842  
Db 301 ACACACACATCGCGCCCTTAGAGAGCGCAGCCCATCTGAGCAAGAGAAAGTAAACAGA 360  
Qy 843 AAGGACACCTCTCTAGGGTCCCAGGGGGCCCGGGCTGGGGCTGCGTAGGTGAAAGGC 902  
Db 361 AAGGACACCTCTCTAGGGTCCCAGGGGGCCCGGGCTGGGGCTGCGTAGGTGAAAGGC 420  
Qy 903 AGAACACTCCGGGCTTCTTAGAAGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGCAT 962  
Db 421 AGAACACTCCGGGCTTCTTAGAAGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGCAT 480  
Qy 963 CGTGTGGCCCTCCCTCCACCTCCCTGTGTATAATATTACATGTGATGTCTGGTCTG 1022  
Db 481 CGTGTGGCCCTCCCTCCACCTCCCTGTGTATAATATTACATGTGATGTCTGGTCTG 540  
Qy 1023 AATGCACAAGCTAAGAGAGCTTGCAAAAAAAAAAAAA 1061  
Db 541 AATGCACAAGCTAAGAGAGCTTGCAAAAAAAAAAAAA 579

Search completed: May 26, 2004, 20:25:25  
Job time : 4421 secs